I presented my work by poster at the Naturalistic Decision-Making Conference in 2017 at the University of Bath as suggested by my co-supervisor at the time Professor Paul Ward. The article I wrote was added to the conference proceedings. Although the research was regarding my PhD I have formed new work since this and have not reported the same data in this thesis. https://www.bath.ac.uk/publications/naturalistic-decision-making-and-uncertainty/attachments/ndm13-conference-proceedings.pdf. ISBN 978-0-86197-194-7

Can Expert-level Cognition be Rapidly Acquired? The Effect of a Human Factors-based Virtual Reality Trainer on Non-Technical Skills in the Operating Theatre

Matthew Pears

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Main Supervisor

Dr David Peebles, BSc, MSc, PhD AFBPsS FHEA

Co-Supervisors

Dr Gurjog Bagri, BSc, MSc, M.Res, PhD AFBPsS

Professor Ashraf Ayoub, PhD, MDS, FDS RCS(Ed), FDS (OS) RCS, FDS RCPS

Professor Minhua Eunice Ma, BA, MA, MSc, PhD, PGCert, FHEA

Professor Paul Ward MSc, PhD, CErgHF, FIEHF, CPsychol, AFBPsS, CSci

Acknowledgements

My family, friends, and colleagues provided me with the material to build the ladder needed to reach this stage. Acknowledging their valuable contribution to my education and to this thesis is the least that I can do:

Dr David Peebles, Dr Gurjog Bagri, Dr Jorge Boehringer, Dr Shawni Dunne, Professor Ashraf Ayoub, Professor Paul Ward, Dr Derrol Kola-Palmer, Dr Yeshwanth Pulijala, Professor Minhua Eunice Ma, Sarah King, Louis Hughes, Chelsea Leadley, Samiya Hussain, Dr Faisal Mushtaq, Sarah Pearson, Adam Robson, Peter Cruickshank, Dr Simon Goodson, Dr David Tyfa.

...and of course, my parents and family who have shown unconditional support throughout.

The real achievement was forming and strengthening the relationships with the people I have met along the way. Thank you all!

Contents

i List of Tables

ii List of Figures

iii Abstract

Background

Restrictions to real-life experiences in surgical training can hinder skill acquisition. Factors such as large student-to-teacher ratios, equipment limitations, or pandemics can reduce access to expert cognition and pedagogical guidance that is required by novices. Additionally, high quality pedagogy from workshops, lectures, and bootcamps are not accessible enough and cannot be attended during pandemic restrictions. Therefore, Non-Technical Skills (NTS) of Operating Theatre (OT) teams need more training content that can provide simulations for training purposes. Patient safety and undesired event prevention can be improved by a scenario-driven approach that is built upon practice and feedback to scaffold cognitive skills for OT trainees. An NTS virtual reality training tool was created and compared to existing theory-based content.

Method

Eighty-two undergraduate surgical students were asked different scenarios and showed decision-making is not distinct as a factor of course year, to generally concur with previous findings. A Task Analysis of a surgical procedure and the OT environment was formed, and 3 experts in surgery were interviewed with thematic analysis of data. The design and creation of the virtual reality instrument then occurred with 360-degrees OT videos. Then, a two groups comparison of a one-hour session with before, during, and after intervention measures compared 14 3rd year operating theatre practitioners. Verbal Protocol Analysis (VPA) of the trainees' sessions were paired with Situation Awareness Global Assessment Technique (SAGAT) scores and rankings for a written decision-making scenario. Post- session reflections were analysed using Interpretative Phenomenological Analysis (IPA) to understand how they experienced the materials and common occurrences between participants.

Results

Thematic analysis of expert interviews revealed rich mental models, tacit knowledge, and purposeful augmentation of NTS as a countermeasure when teaching. This allowed insight into what non-technical elements were feasible when incorporated into a headset. The main VPA findings from the 14 OT trainees suggested significant increase of verbalization around Teamwork and Communication ($p=0.028$). Within this NTS category, significantly more verbalizations for shared mental models for the experimental condition occurred ($p=0.018$). Additionally, a significant increase in transformation of cue meaning to improve understanding of the environment occurred, compared to control condition ($p=0.02$). However, SAGAT scores showed no significant differences in 23 questions for both conditions, this may be a limit in both conditions' presentation delivery as items in the videos are difficult to identify.

Conclusions

Significant results in specific and not all areas highlight complexities in NTS training but is a step towards improved support for OT staff to improve awareness and safety during surgery. Although supposed homogenous technical skills, large variations in participants' decision-making strategies and perceptions of cues may have confounded the intervention effects. During intervention, the control condition used past experiences

to contextually interpret theory to strengthen their schemata in more concrete rather than abstract forms. Real-life scenarios in the experimental condition reduced this need therefore applied their feedback to actual events shown, which may increase transfer of skill to real-life. More sessions over a longer period could observe stronger improvements in the same directions in the current results. Overall, the intervention was equal to or greater than the control condition promoting further research on a greater timeframe and audience.

Chapter 1 - Patient Safety and Surgical Training: Current Issues and Solutions

1.0 Current Issues in Surgical Training

In 2015/16 the National Health Service England (NHS) spent £1.67 billion on compensating patients who had experienced injury, in addition to and legal representative fees (NHSLA, 2016). The NHS Litigation Authority indicated that patient care has been compromised to some extent, due to the costs of clinical negligence claims. Significant increases were predicted (NHSLA, 2016) by the Medical Protection Society (Medical Protection Society, 2018) of reaching £2.6 billion before 2021. The basis was due to a 72% increase in costs in 4 years. Indeed, patients who have undergone surgical procedures find risk of injury can commonly occur during the procedure itself (Vincent, Neale, & Woloshynowych, 2001a, 2001b). More errors have occurred in the Operating Theatre (OT) than any other department (Stelfox, Palmisani, Scurlock, Orav, & Bates, 2006; see also Kohn, Corrigan, & Donaldson, 2000). For example, in cardiac surgeries Young and O'Regan (2010) established a 1 in 300 chance of death due to surgical site infection if admitted to a UK hospital. Approximately two decades before the millennium at least 50% of adverse hospital events occurred within surgical patients (Thomas et al, 2000; Gawande, Thomas, Zinner, & Brennan, 1999). Post-millennium, 14% of surgical patients had experienced an adverse event in some form (Anderson, Davis, Hanna, & Vincent, 2013), and 37.9% of these surgical errors were considered preventable.

Being the main influence in the annual increase of financial disbursement, increase in patient numbers has had a linear growth for more than a decade. There were 16.25 million finished admission episodes in 2015/2016 in the United Kingdom (UK). Yet, the 2017/2018 finished admission episodes statistics show a 2.3% increase from 2015/2016 (digital.nhs.uk, 2018). Certainly, annual admissions for the last decade increase approximately 1-3%. These changes demonstrated that over several years the admissions rate may become substantial and compound admission related issues. If viewed over a decade, compared to 12.68 million hospital finished admission episodes in 2005/2006 there were 28% more in 2015/2016, being 16.25 million. The Royal College of Surgeons (rcseng.ac.uk, 2019) concurrently registered a 27% increase in surgical procedures admissions between 2003-2013.

If only small increases of admissions occur per year, this representative increase over time can amount to markedly more surgical procedures if an approximately linear growth is assumed. This has indicated that the number of surgery related personnel should also increase to ensure a consistent procedure-to-surgeon ratio. For example, Mizuno et al (2014) stated a loss of surgeons over a decade in Japan, accumulating to almost 15% loss in trainee admissions. For the rate of change in the UK, an atypical procedure-to-surgeon ratio may also have developed.

For example, the England Health and Social Care Information Centre 2016 census data for the NHS workforce stated the surgical training group in 2009 totalled 5234 members. They consisted of core medical and core dental training students, along with foundation years 1 and 2 doctors in training across the ten most common speciality areas. However, in 2014 the same group totalled 5207, and 5087 members in 2015. Although the overall NHS professionally qualified clinical workforce has been increasing approximately 1-3% from September 2004 (digital.nhs.uk, 2019) there has been less consistency in trainee numbers. This may indicate the unpredictability for future numbers in both admissions and surgical personnel.

The issues described above were negative outcomes caused by a plethora of factors affecting surgical development and performance. One such tangible influence affecting surgical trainees was the increase of OT time constraint placed upon them. Less access to conventional training by attending live OT procedures led to a reduction in skill acquisition (Alaker, Wynn, & Arulampalam, 2016; Blackmur, Clement, Brady, & Oliver, 2013), with possibly higher workload for surgical mentors, decreasing deliberate teaching time. Consequentially, the occurrence of practice with an array of complex circumstances had likely reduced. Surgical trainees overall have less essential practice to progress to higher levels of expertise, subsequent from increased deliberated training hours (Schaverien, 2010).

Indeed, Caro (1988) defined training as the process of instruction, practice, measurement, and feedback in a task where previous skill proficiency was not at an attainable level prior to training. If there is insufficient chance of experiencing each of these processes, the necessary modification of cognition and behaviour cannot occur. Based upon this, compensatory training when not in surgery is a necessity for societies as the American College of Surgeons or Accreditation Council for Graduate Medical Education (Aim, Lonjon, Hannouche, & Nizard, 2016). This highlighted the limitations of traditional OT training methods and the solution of simulated teaching to compensate.

The effects of a decreasing timeframe for OT practice has been exacerbated by the amplified cognitive demand during training stages places on students. For example, the demand from surgeons, nurses, and physicians for innovative and complex technology has caused a rise to usage of innovative surgical equipment (Banerjee & Gavade, 2015). This was predicted to accelerate surgical equipment value from $10.5 billion in 2016 to approximately $20.3 billion US dollars in 2025 (Grandviewresearch, 2019). If robotic surgery implementation progresses at the rate in previous years Feussner (2017) forecasted surgical robots would account for $20 billion in annual revenue by 2021.

The inclusion of more sophisticated technology can generate a greater learning curve for students than training programmes in previous years. High levels of cognitive workload can delay processing of incoming information therefore slow down the understanding of the information and decision-making (Andersen, Mikkelsen, Konge, Caye-Thomasen, & Sorensen, 2016). The introduction of complex equipment has driven a need to strategize how to reduce cognitive workload of surgeons (Banerjee & Gavade, 2015). This was due to the counter-intuitive nature to add more equipment into a complex system. Although the ability to use more advanced equipment in the OT can reduce workload thus improve surgical performance, more effort is required to use these tools prior to operation. Current teaching methods should be investigated to account for the cognitive workloads of trainees (such as Dias, Ngo-Howard, Boskovski, Zenati, & Yule, 2018) each year, by suggesting how training processes could be adapted to minimize the common occurrences of high cognitive workloads (see Stefanidis, Wang, Korndorffer, Dunne, & Scott, 2010).

Furthermore, one major consequence of the previously mentioned issues in surgical training has been a negative learning experience for the individual. Fewer OT opportunities have increased the workload on trainees when in the OT, with higher learning curves causing synergistic and iterative adverse effects. Approximately 43.1% of surgical trainees in the United Kingdom have experienced negative events. These involve unrealistic knowledge expectations, and due to time constraints may receive poor or inadequate pedagogy (Chapman et al, 2013). Insufficient surgical education and collective negative effects could be a factor in the surgical residents' figures as previously discussed (Nakayama et al, 2016). This may also result in demotivation, with motivation one of the most important factors in the success of training (Chapman et al, 2013). Hence, it may be suggested that the limited hours of vital OT practice combined with higher workload have substantially affected students' learning experiences and provided evidence for the revision of current skill acquisition methods.

1.1 Current Solutions to Surgical Training Issues

There are different types of information in the OT that a team call upon to make intraoperative decisions. These are patient information, medical knowledge, surgical procedure knowledge, and visual and tactile information (Vannaprathip et al., 2016). Physical skills such as good hand-eye coordination, manual dexterity, and accurate psychomotor skills have also been categorised as 'Technical Skills' of the individual (Moorthy, Munz, Sarker, & Darzi, 2003). The principal elements in current simulations to support training are primarily comprised and oriented around these technical proficiencies (e.g. Mercier, Chagnon-Monarque, Lavigne, & Ayad, 2018). This has been beneficial as technical skills can be quantified, i.e. knots tied per minute. Strong technical skills have also been correlated with better operational outcomes in general.

These information types can be categorized by Rasmussen's model of human behaviour (Rasmussen, 1983; 1986; 1990) which include knowledge, rule, or skill-based information. Knowledge-based information includes but is not limited to, interactive text displays linked to tools, equipment, and body parts, which describe the items of interest. Rule-based information includes but is not limited to, the systematic procedural steps that occur during the specific operation. Skill-based information includes physical practice and kinaesthetic feedback of the tasks. With the large range of information and skills required, most institutions focus on improving these technical elements.

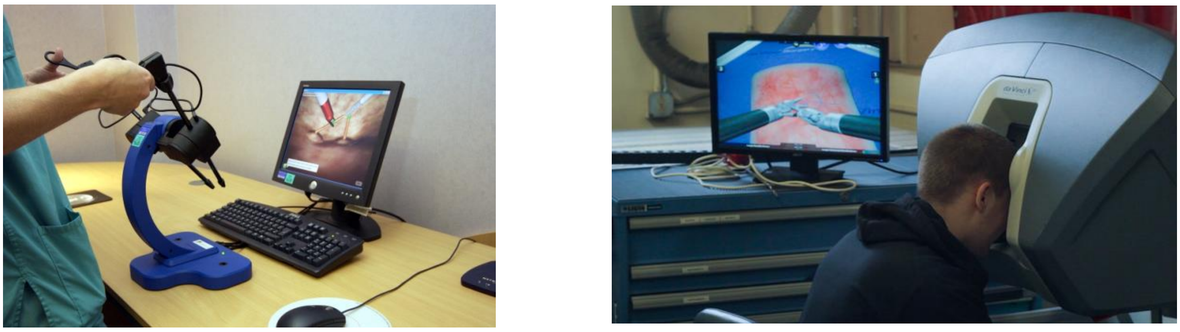
For this training, one prevailing method has been the use of simulation and this has included one or more types of information, dependent upon the type of technical skills it aims to improve. Simulation can involve human and system interaction that can provide means of effective, cost efficient, and safety-orientated human training, analysis, and evaluation (Hancock, Vincenzi, Wise, & Mouloua, 2008). Through replication and duplication of all or specific parts of a system, an interactive simulator may be created. Simulation of a real-world environment overcomes the limitations of training in a real-world system and can break down complex medical and surgical elements into sub-components and allow modular learning for students.

With attention to technical improvement for surgical trainees, use of simulation has been the primary technique used to enhance trainees' skills during operation. With most simulation comprised of skills- based information, they provide a variety of options to develop skills by using inanimate (such as box trainers), ex vivo, and in vivo animal models, digital simulations, or fully simulated OT rooms. These methods provide opportunities to improve preparedness for the OR through a range of fidelity/realism. They have allowed the improvement of skills without compromising patient safety (Torkington, Smith, Rees, & Darzi, 2001a; 2001b). Importantly, as learning in modern OT environments requires greater rates of adaption due to their complexities, novices may experience large learning curves (Banerjee & Gavade, 2015). These large curves can be reduced by simulations as they provide students with techniques to manage with complex tasks, with transferability to theatre (Ahmed, Aydin, Dasgupta, Khan, & McCabe 2015). Simulations also offer standardized scenarios to ensure consistent and standardized procedures and feedback, along with facilitation to improve skills/techniques that were highlighted as suboptimal. For example, simulation as an analytical tool for performance in students allows systematic evaluation of their hands-on and cognition related progression (Berridge, Jain, & Shekhar, 2019). Moreover, comparisons within large amounts of data can safeguard standardized skill levels in the same cohort.

For example, technical skill training investigated by Kang et al (2015) used a state-of-the-art high-fidelity simulator (Mimic dV-trainer, 2015) for training to use the latest 3D minimally invasive robotics tool named the 'da Vinci surgical system' (davincisurgery, 2020, see Figure 1). They assessed 30 medical students, who received 4 hours of Mimic dV training over different time course. Kang et al (2015) concluded daily one-hour practice sessions resulted in the most improved scores, in terms of time and repetitions needed to complete a simulator task. The advantages of the quantifiable and repeatable nature of technical simulation were exemplified in their study.

Further pragmatic evidence of the effectiveness of simulation for technical proficiency improvement has been established in several regions in the UK but mostly within Yorkshire and Humberside. Simulators were trailed with focus on performance-driven evaluation of students (Yorksandhumberdeanery, 2016). This included a 4-day urology boot camp with use of actors/simulated rooms that allow staff to observe and give feedback to students (Berridge, Jain, & Shekhar, 2019; Medical Education Leeds, 2019; Pears et al, 2021). These benefits have provided initial rationale for the enhancement of current simulators, as they may be a viable method to address the previous limitations.

Figure 1: Digital Surgical Simulators - Mimic dv-trainer (right) and laparoscopic trainer (left) used to train students with a variety of procedural-, knowledge-, and skill- based programs.



1.2 Future Suitability of Simulations

Through this viewpoint, a rationale for the progression of surgical simulations has been present in the literature. Yet other factors such as perceptions, implementation, and uptake success should also be evaluated to understand the likelihood of future use. Future implementation of the full range of simulators may not be supported in the literature. There may be risk of some of the tools becoming obsolete or replaced by a more efficient training tool for select skill acquisition. Assessing the differing types of simulation Yiasemidou, Glassman, Tomlinson, Song, and Gough (2016) investigated the prospect of surgical simulation usage from experts' perspectives. Their aim was to evaluate experts' opinions towards current and future roles of education utilizing surgical simulation. Four types of simulation were compared: cadaver, animal models, synthetic models, and virtual simulators. They analysed 36 questionnaires from 15 Heads of Schools of Surgery, 21 local education and training curriculum members and gave semi-structured interviews to five surgeons with national educational roles.

Thematic analysis of the semi-structured conversations and statistical analysis of questionnaires were produced. The results suggested that simulation was accepted as not only useful to novices for basic skill acquisition, but also for experts as they could also use simulation to learn complex procedures. Participants largely agreed with the notion of formal accreditation of simulation centres (60% agreed 20% strongly agreed). When asked if trainees should receive compulsory simulation training, prior to performing a procedure for the first time, 53.33% strongly agreed/agreed with this statement. To highlight a key report mentioned by Yiasemidou et al (2016), trainees were in favour of the inclusion of simulation in their curriculum when delivered through training centres (asit.org, 2011). Therefore, the results by Yiasemidou et al (2016) also subsequently support the use of simulation for skill acquisition.

However, deciding which type of simulator most suitable to develop has been less apparent. More specific outcomes by Yiasemidou et al (2016) suggested differing opinions when participants were asked to order 4 typical simulators in terms of preferred training method. Somewhat predictably, cadavers were first choice by 70% of all participant. Also, animal models were second choice as 44.8% of all participants agreed. When comparing opinions between votes, virtual simulators had relatively less verdict. *Between* each type, virtual simulation was concluded as a tertiary training choice, as 41.45% of all participants opinionated the notion. Yet as a *within* simulator types, virtual simulation exceeded the percentages of votes for animal models and synthetic models combined in the 'first option', therefore being a first choice after cadaver usage, as the primary format of training. This showed that although virtual simulation, such as a LapSim (surgicalscience.com, 2016) was a tertiary option, therein lies the potential for it to be a primary choice under certain circumstances as selected by the training program directors who participated in the study.

To emphasise the importance of this, when comparing the figures with a standard distribution (probability density function) the frequency for response for cadavers should have approximately a 68% vote distribution, the frequency to animal models (being 2nd) should approximate the 27%-mark, virtual simulation (being 3rd) should have a 4.7% vote distribution, and synthetic models (being last choice) with 0.3% of votes. Cadaver votes had a distribution of 68.1%- this result as first choice was relatively appropriate being much of the probability density. Yet virtual simulation received 16.67% of the remaining votes and overtook the choice of animal models (11.11%). Further research could gain extra insight into how and why the prospective utility of virtual simulation had more popularity than other existing simulations.

1.3 Virtual Simulation

The traditional equipment for virtual simulation had included mediums such as PC/TV monitors, CAVE (Cave Automatic Virtual Environment) systems, and digital projectors. This provided a ranging low-high fidelity spectrum which virtual surgical simulators can be delivered upon. Typically, dummy surgical instruments represent the interface with the surgeon. The last decade has developed many devices both with force feedback (haptic feedback enabling kinaesthetic perception), and without. These have been interactive and provide the user with a sense of presence and engagement in a computer-generated Virtual Environment (VE) (Repetto & Riva, 2011; Bordnick, Traylor, Carter, & Graap, 2012). This VE is a fictive and secure environment as it has no physical threats. As it is virtual, the user can alter any aspect of the environment. Being a virtual composition, it aims to represent and correspond to reality. Consequently, it is an intermediating platform as it sits between the perceived entities in the real world, and the safety of practice in a virtual world (Aebersold, Rasmussen, & Mulrenin, 2020). Furthermore, recent virtual surgical simulators utilise advances in software rendering to enable systematic task processes. Therefore, they offer the opportunity to increment procedural task complexities that come with modern techniques and consequentially trainees can perform procedures that are more elaborate.

Considering properties of the healthcare workforce which could increase virtual simulation, the compatibility with the current student generation should be considered. These individuals have been theorised to have increased technologically orientated abilities and critical appraisal of tasks. To expand on this point, the millennial generation (aka Generation Y) referred to individuals born from 1982 to 2004 (DiLullo, McGee, & Kriebel, 2011). Prensky (2001) coined these individuals as "digital natives" as a general description for their native abilities in digital language including the internet, computers, and video games. Prensky (2001) suggested that Generation Y has a different way of processing information due to the effect of technology usage from a young age. Previous teaching methods used to teach Generation X are sub-optimal. Seymour (2008) also suggested new surgeons benefited from simulated virtual training which resulted in less OT usage and improved technical performance (Nagendran, Gurusamy, Aggarwal, Loizidou, & Davidson, 2013).

With real-world application of these devices Nagendran, Gurusamy, Aggarwal, Loizidou, and Davidson (2013) suggested that laparoscopic surgery outcomes improved specifically due to virtual training. This is in terms of significantly decreased operating times and technical performance, compared to no training or the standard procedure of box training (2 laparoscopes protruding from a box that trainees can practice with). Concurrently, Piromchai, Avery, Laopaiboon, Kennedy, and O'Leary (2015) evaluated the efficacy of virtual surgical training in terms of patient outcome within the field of Ear, Nose, and Throat (ENT) surgery. Their objective was to assess whether current evidence supported virtual integration of surgical simulation into training programmes with the additional aim that skill outcome would equate to, or be better than, skills resultant from conventional training methods. Participants showed a range of skill levels, from medical students through to more advanced surgeons in training. They concluded that surgical virtual training significantly improved the psychomotor skills and completion times for trainees in theatre, therefore improving technical skills. Additionally, medical students who received virtual training in a controlled environment showed significantly higher procedural scores and anatomical identification.

Piromchai, Avery, Laopaiboon, Kennedy, and O'Leary (2015) suggested that this available evidence supported the notion that technical skills can be equally learned, or better, via virtual simulation-based training. This was in comparison to conventional training methods, and irrespective of the application being within an OT, or a controlled environment. They suggested that virtual simulation could be a supplementary training tool in its current state, but not as a replacement for already formally implemented surgical training activities. Following this, Alaker, Wynn, and Arulampalam (2016) meta-analysed 31 randomised controlled studies comparing virtual laparoscopic simulation to alternative models of training. Their results suggested proficiency-based simulation was significantly more effective than video trainers. Importantly, VR simulation was at least as effective as box-trainers. From this, they suggested that there was a necessity to incorporate such virtual training into surgical curricula.

Illustrating the array of virtual reality simulator applications, implementation in psychological therapies has had comparable success to aviation. Their effectiveness in psychological therapies has shown to be equivalent in effect to the gold standard of in-vivo exposure (Meyerbroeker, Morina, Kerhof, & Emmelkamp, 2013). Virtual simulations in psychological therapies have been developed to treat affective (Opris, 2012), cognitive (Parsons, 2009), and specific cognitive components such as attention (Aman, 2020), memory (Moffat, 2009), and executive functions (Armstrong et al, 2013). This has meant treatments for anxiety disorders; specific phobias (Shiban, Reichenberger, Neumann, & Muhlberger, 2015), post-traumatic stress disorder, and panic disorder have been improved due to the advantages and applicability of virtual simulation (Diemer, Mühlberger, Pauli, & Zwanzger, 2014; Glotzbach-Schoon et al, 2013).

1.4 Virtual Reality Headsets

From this overview of simulation in healthcare, within the range of virtual training there has been one method of technical skills teaching has been attempted via virtual reality headsets. For decades, the advanced technological criteria to support their theoretical principals and design have not been met and results have been mixed. However, due to the propitious advancements in different areas of technology, such as nanotechnology, VR headsets have gained the required technology. VR headsets are display devices worn on the head of a user, with high-definition optical displays positioned in front of each eye to create binocular vision. These optical displays present 3-dimensional structures/images/videos, producing stereoscopic vision. In addition, VR headsets have included 360 degrees (360°) audio positioning, head positioning, and geometric head detection systems, which track head movements and can instantaneously represent these movements in the user's 360° environment (Oculus.com, 2020). Marking the start of more advanced VR headsets, the release of the Oculus Rift Development Kit 1 in 2013 (Oculus, 2020) promoted the increase in global interest and rapid release of other headsets by a plethora of companies. By tracking the technology being bought and globally shipped, the International Data Corporation (IDC, 2019) forecasted shipment for VR headsets to go from 7.2 million units in 2018 and 8.9 million units in 2019, to 36.7 million units in 2023.The Oculus Rift Development Kit 2, Oculus Rift Consumer Version 1, and the HTC Vive have been the most use in academic research and in consumer usage.

1.41 Immersion and Presence

VR headsets can provoke an increased sense of presence (Waterworth, Waterworth, Riva, & Mantovani, 2015) when compared to widely used displays (however see Kim, Rosenthal, Zielinski, & Brady, 2012). Additionally, they provide a significant increase in presence to previously released HMD's as technology had evolved to better address previous issues (Cummings & Bailenson, 2016; Lin, Park, Liebert, & Lau, 2015). Headsets can provide the user with an egocentric view perspective (see Alsmith, 2015), which has been suggested to support a more natural mapping of the user into their VR (Hartmann et al, 2015). These headsets create strong sensory richness, and similar to reality the first person perspective facilitated embodiment in the VR world. This high amount of feedback facilitates a feeling of presence in a first-person VE (Sundar, Go, Kim, & Zhang, 2015). This increased facilitation of affective experience can provide the user with increased cognitive engagement- which cyclically enhances sensation of being within the VE.

The underpinning mechanisms of presence are multifaceted. The principal components of immersion and flow have relevance in the design of the tool to help users achieve a feeling of presence. With increased sophistication of simulation (representing higher immersion) increased presence can more readliy occur (Cadet & Chainey, 2020; Hassan, Julha, Sjoblom, & Hamari, 2020). However, this is under the conditions that flow also meets certain demand characteristics to facilitate higher levels of immersion. These factors aid the user to teleport themselves (presence and telepresence are used synonymously in VR literature). Flow loosely translates to the utility of a user's senses during VR use. A 'good' flow state represents intrinsically optimised use of attention and a natural feeling when performing tasks in the system.

Flow has been measured by the Flow State Short Scale (Jackson, Martin, & Eklund 2008) which consists of nine main theorised aspects of flow: Challenge and skill balance, loss of self-consciousness, sense of control, clear goals, unambiguous feedback, concentration on the task at hand, action-awareness merging, time transformation, and autotelic experience. Longer flow measures have been developed however were not practical and the shorter 9-item version provided valid and reliable measure of a user's global flow construct.

There are commonly occurring methods/materials in VR usage and have been shown to provide suitable levels of immersive experience (Pallavicini, Pepe, & Minissi, 2019) and the resources in many cited studies in this project allowed an adequate amount of presence and flow to be integrated into the system, by means of high-quality video, intuitive and responsive hardware, and simplistic and natural interaction methods. The consensus is that flow and immersion obtained using common VR setups is sufficient to not confound studies, without being too resourceful (ie adding haptic touches). When considering the goal of creating an impactful pedagogical intervention towards increasing understanding of Operating Theatre processes, increasing presence beyond a typical level may detract from this key investigation (i.e., adding stereoscopic depth for increased spatial presence can cause discomfort/headaches). Additionally, the learning outcomes in studies similar to the current project are attributed to an array of factors that compile to create the final intervention. There would be a severe increase in resources/time to be able to attribute each immersive property. However, by utilising a common setup (monoscopic videos, standard 90 to 120-degree field of view, 30 frames per second) the global output of the intervention has a suitable level of flow and immersion for users and concurs with most contemporary research practice (Parong et al, 2020; Weech, Kenny, & Barnett-Cowan, 2019).

1.42 VR headsets in Healthcare

Moving to comparisons against VR headsets, Botden, Torab, Buzink, and Jakimowicz, (2008) suggested there were no significant improvements between the use of box-trainers and VR technology.. However, other research prior to VR headsets such as the Oculus Rift, has been shown to be equal or better than surgical box-trainers (Alaker, Wynn, & Arulampalam, 2016). This contradiction in findings and many other comparative research outcomes may be due a pivotal moment in the effectiveness of the VR headset used in each study. There may need to be replication with usage of the new generation of VR headsets which have excelled far greater than previous versions. At the release of the new generation of headsets there was support from several domains in modifying cognition and behaviour (Aiken & Berry, 2015; Gahm, Reger, Ingram, Reger, & Rizzo, 2015; Koo et al, 2015; Nori et al, 2015; Shaw et al, 2015). A surge of research had emerged in academia using this advanced equipment however there was a necessity for more OT-based training investigations.

Pragmatic support of VR headset inclusion in high-quality surgical training and assessment also appeared from a concurrence of medical technologies diverging to 3D display forms. For example, previous laparoscopic surgery required specialist-training procedures to overcome 2D visualizations of the 3D surgical field (Bahrami et al, 2014). However, 3D laparoscopic surgery has gained popularity, partially due to the reduction of the cognitive workload of the surgeon. Cologne et al (2015) compared 2D and 3D laparoscopy effects on the learning curves of novice medical students. With the single change from 2D to 3D visualization PEG (basic laparoscopic training technique) transfer completion time fell by 45.5%. Mean PEG drop rates were zero, and errors halved compared to the 2D condition. These results similarly occurred in a suturing task performed by novices, which also provided evidence for laparoscopic training in 3D. Thus, leading robotic technologies such as the D*a Vinci Surgical System* (davincisurgery, 2020) have integrated stereoscopic cameras and built in 3D monitors to provide the advantages of depth perception to the surgeon. Also, comparably simpler laparoscopes such as the Endoeye Flex 3D (Olympusamerica.com, 2020), have significant advantages over previous 2D displays and have begun increasing 3D usage. Yet, very few studies exist that identify the value of such features of VR headset in surgical training, amongst other theorised benefits. Therefore, it was necessary to explore the reliability, proficiency, and educational value of this developing technology.

2.0 Critique of Relevant Contemporary Research

To explore the educational impact of using VR headsets with technical skills training, a search of literature was performed from the University of Huddersfield's Summon (2019) search engine, Google Scholar (2019), and Ovid (2019). Two factors of interest were the criteria required: for potential studies to involve technical training for novice OT professional, and to use an advanced VR Headset. One PhD study was found and met these criteria at the time of searching, being late 2016. This PhD study was in the process of creating and assessing a VR surgical training tool for trainees in Oral and Maxillofacial Surgery (OMFS). After contacting the lead researcher of this study, a collaboration was formed at the University of Huddersfield. This collaboration was to evaluate the VR technical skills tool created by Pulijala. Its purpose was to understand the advantages and disadvantages of this training setup, and explore if it could improve patient safety via enhanced surgical training. If this method would be the best available to address the described issues, then the insight on how to develop this combination could be further investigated.

The collaborative effort was published in the Journal of Oral and Maxillofacial Surgery (see Pulijala, Ma, Pears, Peebles, & Ayoub, 2018). The following section describes the initial point of collaboration which was the basis of co-authorship and influenced consequent research. There were other areas within Pulijala, Ma, Pears, Peebles, and Ayoub (2018) which the researcher of this project was not specifically involved in (*see the article for full details).* The researcher of this project had direct influence/collaboration/involvement with elements in the section below and has been expanded in subsequent chapters.

The VR technical skills tool was nearing completion when the collaboration was formed. It was built with a menu driven structure therefore was minimal complexity and minimal change of being lost in navigation to a desired location. The VR layout was a supportive structure for users to advance their understandings in a logic sequence when they felt comfortable in doing so. It was compared to current educational methods within the area of Oral and Maxillofacial Surgery. The content consisted of the traditional media associated with surgical training; the tools of the surgical procedures, relevant human anatomy, and an overview of the procedural steps. The final VR tool consisted of 2D surgical videos, and 3D computerized model 'assets' whereby users could zoom, rotate, and look inside each asset presented. These assets were skulls, arteries, and veins within the head, real-life scans of patients face and teeth, and the tools used in most common Oral and Maxillofacial surgeries. Therefore, all content was considered as technical material/information.

The VR tool was created for use by OMFS surgeons either undergraduate or postgraduates in the UK. However, in a universal sense the VR tool was able to be implemented in most institutions to trainees that were not fully independent in performing in OMFS. This meant that although UK surgical trainees could have been sourced, the availability of trainees in locations where there was a higher number of students would enable easier access and increase the number of overall participants. Based on this, both public and private universities and hospitals located in India were selected, as the ratio of trainees to surgeons was higher than in the UK. A control condition of 45 participants and an experimental (VR) condition of 51 participants were compared and was a single-blind, parallel, randomized trial. Questionnaires were given to participants before delivering these conditions, which measured their previous usage of technology, training materials used, confidence, and knowledge of equipment, anatomy, surgical steps, and other factors. The intervention stages were given to participants for the same length of time, with material which differs in how it was presented but had the same underpinning information embedded within each condition. Post-intervention measures were the same as the pre-intervention measures and performed 5 minutes after intervention.

The results displayed a significant increase in perceived self confidence levels ($p<0.001$) in both conditions after the intervention. Comparing relative improvement in confidence, the experimental condition effects significantly increased scores compared to those in the control condition ($p=0.034$). The between subjects' results indicated a significant increase in first year resident's confidence compared to third year residents ($p=0.001$). In terms of knowledge, initial ANOVA showed the experimental group ($p= 0.024$) and the control group ($p= 0.025$) significantly improved post- assessment measures/scores compared with the same measures before intervention. However, the experimental group did not significantly increase scores. Therefore, the experimental condition did not significantly increase knowledge compared to the control condition.

2.1 Pulijala et al (2018) Review

Pulijala, Ma, Pears, Peebles, and Ayoub (2018) was the most relevant research output at the time of searching which investigated how surgical trainees may be helped with virtual training outside of the OT. The content, hardware, design, and user interaction combined to form both benefits and difficulties in information presentation, and therefore value of information for the user. The results obtained directly influenced the subsequent direction taken in the remainder of this current project. The critical discussion of this study highlighted several substantial advantages and disadvantages, which were taken into consideration.

The results in Pulijala, Ma, Pears, Peebles, and Ayoub (2018) described all participants experiencing increasing perceived confidence in their technical skills from their respective interventions. As previously mentioned, the negative experiences and learning curves of training may be factors which reduce confidence in the OT. Therefore, any training material which can deliver the required information with little or no negative feedback and allows the user to control their pace of incoming information, may counter these issues and aid in boosting confidence.

However, the experimental VR condition was significantly more effective for 1st year trainees compared to 2nd and 3rd year students for increasing perceived confidence. This meant there were factors in the intervention which caused this difference, for only the experimental intervention. With participants being separated by their year in course education, one likely difference would be the type of knowledge presented to each group. The optimal effectiveness towards boosting 1st year students' confidence can suggest the type of information presented was more appropriate to them. Participants in 2nd and 3rd years may have more knowledge of the content therefore not instigating a boost in confidence as their understanding remained unchanged. If information was too basic/simple for the 2nd and 3rd year students, the significant increase in outcome measures for the 1st year participants may only be due to the novelty of the information they were presented with. This may have caused increased confidence from understanding this new information.

This notion cannot be true however, as this difference between 1st year and the other two years did not occur in the control condition, although the material presented was the same. Additionally, knowledge did not significantly change, both within and between conditions. Therefore, the significant increase in 1st year confidence must have been from the VR delivery method, more than from any increase in knowledge from new material presented to them. This significant confidence increase found for the 1st year group compared to 2nd and 3rd years in the experimental condition can be reduced to the novelty of the VR equipment, but it is unclear the mechanisms which were deployed from to its usage. There may be individual differences in the 1st year VR condition which explain their changes.

Moving from *why* confidence increased towards *how* it can form real-world consequences – confidence can produce negative effects when considering that technical knowledge did not also significantly improve. In Pulijala, Ma, Pears, Peebles, and Ayoub (2018), there was not a significant increase in knowledge in the experimental, compared to the control, conditions. Certainly, simulations have been used to improve confidence in undergraduate trainees (Bommer et al, 2018). Yet only this component was increased which formed disparity between perception of skill and possible real-world performance. Significantly boosted confidence may have a negative consequence if it is not accompanied by the related improvement in performance. The effects of overconfidence are biases that suggest objective accuracy of performance is mismatched with the individual's subjective judgement of the skills relating to their performance. This can form the Dunning- Kruger effect (2011) if individuals are not equipped to understand they may possess a mismatched perspective (Pallier et al, 2002; Borracci & Arribalzaga, 2018). Therefore, to avoid such consequence a measure of the impact of VR technology was also needed for each participant, to account for the possible discrepancy when compared to subjective confidence of performance. This would help researchers understand if virtual reality may augment confidence in a disproportional factor with those that may have a susceptibility to reduced confidence levels (being 1st years), compared to other teaching methods. Further research could establish the best approach to control for consequences of overconfidence from VR, if the paired performance outcome improvement is absent.

Certainly, cognitive measures within Pulijala, Ma, Pears, Peebles, and Ayoub (2018) could have been embedded to capture the processes involved in the changes to participants. The developments in confidence improvement could also have been explored using cognitive measures. In order to improve their skills in surgery, an individual must process and transform information to allow cognitive changes which will produce improved behavioural outputs in the OT. This could be in any skillset i.e. anatomical locations, or procedural steps. Cognitive assessments can surmise the likelihood that information presented would be retained over a longer time period when compared to other training methods. This has importance when the data gathering period of studies may be short, a single-session, and without follow up. For Pulijala, Ma, Pears, Peebles, and Ayoub (2018), the retention of the information presented would need to be stored for use over the long-term yet measures of short-term memory were used (as participants were assesses immediately after intervention). No underpinning and in-depth information was collected on how the intervention influenced the participants abilities to better understand the information they were presented with, even if they have seen the information before. Additionally, there may be a counter-intuitive element to the addition of more technological interactions. With the same perspective, Scafa, Serrani, Papetti, Brunzini, and Germani (2019) stated that with possible information overload there was a need for measurement of cognitive conditions during simulation as to prevent overload of information. The VR tool used in Pulijala, Ma, Pears, Peebles, and Ayoub (2018) had no feedback on the level of users' cognitive states upon usage and may have provided too little or too much information at various points. Cognitive feedback would have helped to understand how the training tool was influencing the users' attention, perception, and type of memory amongst other elements.

Furthermore, cognitive measures of users could help to illuminate any restrictive effects from the type of information presented. Training and assessment in many medical and surgical educational syllabi and curricula only consider one form of appraisal, being technical skill. When compared to the environments of real-life operations, singularly technical skill simulations omit a large amount of rich information that may aid in patient safety. Technical skills teaching has sound rationalisation as the aim of assessment with use of simulation generally endeavours to produce competent specialists. The common premise has been that assessing technical skill predicts the abilities of the individual (Graafland, Schraagen, Boermeester, Bemelman, & Schijen, 2015). The advantages surround evaluation of surgical competency by use of structured assessment scales to objectively measure technical skills (Vassiliou et al, 2005; Reznick, MacRae, 2006). However, this information has limitation in improving users' global surgical ability in relation to patient safety. There are many other important perceptual-cognitive requirements which could be adapted to augment an individual's proficiency within their real-world environment. A technical skills device without cognitive measure and with limit to one type of information, may be a bottleneck and impeding the transformation and improvement of users' overall cognition and behaviour.

With improvement to patient safety and quality of surgical training being the primary goals, the main perspective reviewed has been to focus on surgical skills improvement by orientation around technical proficiency. Traditional 'hands on' training with less access to OT practice has adapted to create technical simulation outside of theatre- with much success. Nevertheless, although technical competency links to overall patient safety improvements, the future layout of complex surgical environments will demand more research into other closely associated factors, especially cognitive training. The ability to possess other perceptual and cognitive skills through simulation will be essential to improve future training experiences and delivery. Yet existing surgical simulators offer little training in cognitive skills (Shiralkar, 2011) associated with higher level mental functions such as decision-making, communication, and problem-solving. Modern complex human–system environments implore training from other perspectives which consider optimising the abilities of the human within the system whilst being designed to be a positive learning experience for trainees. Due to the high amount of information received, there may be more dependency on the abilities of humans within the system. For example, specifically how they can understand and transform data/information for their own purposes, in addition to improvements in technical skills.

Chapter 3- Human Factors Techniques and Non-Technical Skills

3.0 Human Factors

From reviewing the problems in surgical training and understanding the limitations with contemporary attempts to minimise these issues, a need for expansion into cognitive proficiency and more in-depth measures was highlighted. To assess the cognitive challenges, either in the OT or caused by a training simulation, psychological methodology was needed to be able to measure such implicit events. Branching out of cognitive psychology into applied cognitive research, is the approach of Human Factors Psychology. Human Factors methodology (see Stanton et al, 2013) has been used as a tool to obtain insight into the cognitive differences/occurrences between individuals with differing skill levels. Therefore, this had the tools needed to further investigate the role of cognition in the complex system of surgery.

Human Factors integration has examined the inter-relations between the individual, the system, and the task at hand. Through techniques such as cognitive interviews, it can reveal differences in underpinning cognitions in experts and trainees (Flin, Youngson, & Yule, 2015). Its history and implementation in a surgical setting is undeveloped when compared to other domains that have complex dynamic events (Flin & Mitchell, 2008). They provide support for medical and surgical practice to benefit from Human Factor methodology.

Human Factors initially orientated around the adaptation of human interaction with machines or objects. This meant humans were the primary object of focus, and of change (Wickens & Holland, 2000). A significant increase in Human Factors research occurred during World War 1. As new technologies enabled new military equipment (such as tanks and aeroplanes), there was need to rapidly train humans to succeed in their goals. Users would be dependent on the technologies in novel, complex, and dynamic environments (Meister, 1999). Throughout World Wars 1 and 2, psychological studies within military context suggested it was possible to identify and filter select human characteristics. This could determine, for example, if a human would become a successful pilot. However, technological advances outpaced the human ability to adapt to their complex usage. Negative events such as injuries, loss of resources, and deaths increasingly occurred (Fitts & Jones, 1947). From this, the design of planes incorporated the human measure and movement, and became user orientated. Within the automobile industry, human behaviour was also incorporated in the design of new automobiles by assessing natural interactions, movements, and needs of the user (Forbes, 1939). From the 1950's several military based research laboratories developed this definition of Human Factors (see Licht, Polzella, & Boff, 1993) and branched into an ergonomics-based Human Factors approach and emerged to be a separate and explicitly defined field.

However, with increased understanding of cognition in system usage, human limitations became apparent in systems design. Therefore, the definition morphed once more to become primarily cognitively orientated, with less emphasis on the ergonomic perspective. Depending on a researcher's goals and perspective, Human Factors may refer to the interaction of a user and equipment, dynamics of a system including interactions of many individuals, or the cognitive elements of individuals which are affected by, and effect upon, a system (Klein, Moon, & Hoffman, 2006). This project used the latter definition of 'Human Factors', being a perspective orientated around cognitive functions and consequences of system usage.

3.1 Effectiveness of Human Factors Outside of Surgery

Initially, surgical simulation usage originated from the transfer and application of Human Factors research (Salvendy, 2012) from the military and aviation upon the challenges within high-risk surgical settings. The use of simulators in aviation has shown to be in-depth, less expensive, safer, accessible, and produce more effective training opportunities than an actual system may provide (Moroney & Moroney, 1999). It has been widely accepted that pilots reach certain simulation proficiency levels before managing an aircraft. Simulated training was implemented in surgical education for the same primary reasons, as well as a supplementary method to compensate for the reduction of time spent training surgical students inside the OT (Brunckhorst et al, 2015). Although Human Factors in a surgical setting has been many steps behind other high-risk organisations, the understanding of both organisational and Human Factors research in areas such as aviation has guided surgical simulation development to recently refine performance of both system and user, increase error prevention, and virtual environment development (Banerjee & Gavade, 2015). As the reliability of medical and surgical practice could theoretically significantly benefit, there has been support for further promotion of this paradigm shift towards Human Factor methodologies with a similar effect as with aviation (Agha & Fowler, 2015).

In research which had not utilised Human Factors methodology, a plethora of research has supported the implementation of VR simulation in technical skills training for surgery. However, there has been an assumption within surgical literature that technical skills competency can predict many abilities (Graafland, Schraagen, Boermeester, Bemelman, & Schijen, 2015). These skills alone are not adequate to preserve or prevent patient safety, or reduce errors (Flin, O’Connor, & Crichton, 2013). Yet, in research which had utilised Human Factors methodology, a new understanding of the performance of individuals was formed that did not necessarily refer to technical abilities (Flin et al, 2003). There are more skills required from the surgeon that complement technical skills and are commonly referred to as non-technical skills (NTS; Yule, Flin, Paterson- Brown, & Maran, 2006).

3.2 Non-Technical Skills

Research in Healthcare with utility of Human Factors Methodology has helped to uncover NTS. Following the classification system by Flin, Youngson, and Yule (2015) Decision-Making, Situational Awareness, Communication and Teamwork, and Leadership were formed through iterative research. There has been a growing interest of human factors training in surgical education curricula, such as a surgeon’s non-technical abilities, when compared to emphasis and focus on technical skill acquisition (Lin, Park, Liebert, & Lau, 2015). Non-technical skills (NTS) involve cognitive and social skills that are semi-independent from clinical knowledge, dexterity, and use of equipment, yet are linked to, and necessary for, the high skills output. Situational awareness and decision-making are categorised as cognitive elements, with the remainders being socially orientated elements. Being suggested as requirements to achieve competence in the Operating Theatre (Yule, Flin, Paterson- Brown, & Maran, 2006), NTS following appropriate levels of knowledge and technical skill. From understanding the successful implementation of Human Factors in aviation (Flin et al, 2003), oil exploration (Flin, O’Connor, & Mearns, 2002), anaesthesia (Fletcher, McGeorge, Flin, Glavin, & Maran, 2002), and nuclear power, non-technical skills training is becoming more embedded to the same level in surgery (Agha & Fowler, 2015).

There are many basal similarities across such high-risk domains: stress, tiredness, and fatigue (Balch, 2010), communication and leadership issues (Lingard et al, 2008), and ineffective teamwork (Bethune, Sasirekha, Sahu, Cawthorn, & Pullyblank, 2011) are a few. These are common Human Factors comparable in terms of their roles in predictable and preventable errors within the organisations previously mentioned. In surgery, NTS training is important in the prevention of errors; for example, 27% of preventable surgical errors have been suggested to be caused by NTS, being cognitive and diagnostic errors (Wilson, 1999) and 43% due to errors in communication (Anderson, Davis, Hanna, & Vincent, 2013; Gawande, Zinner, Studdert, & Brennan, 2003). Hence, non-physical factors involved in procedures have also caused patient injuries besides surgical techniques (Flin & Mitchell, 2008).

NTS are markers of both behaviour and cognition that provide contribution for the increase or decrease in performance measures (Flin et al, 2003). Thus, their impact on behavioural outcome influence the efficiency of technical skill acquisition. Reciprocally, a lack of technical skill can also decrease the efficacy of non-technical cognition and behaviour. For example, if a surgical procedure was not fully understood by a trainee, then non-technical abilities such as Situational Awareness cannot be developed without this initial technical understanding. This is because the individual has limited procedural knowledge therefore, they cannot expand their prospective non-technical comprehension of how to predict what could occur with their patient.

On systematically reviewing the impact of anaesthesiologist's technical performance and non-technical skills, Hull et al (2012) suggested a correlation between these two skillsets whereby non-technical performance can enhance technical proficiency or if absent can contribute of reduced proficiency. Black, Nestel, Kneebone, and Wolfe (2010) also demonstrated a strong relationship between technical error and teamwork failure in the OR. Indeed, there are many reports that highlight the links between adverse effects on healthcare and non-technical performance markers, beside a lack of technical skills (Fletcher, McGeorge, Glavin, & Maran, 2002; Webb et al, 1993; Cooper, Newbower, Long, & McPeek, 2002).

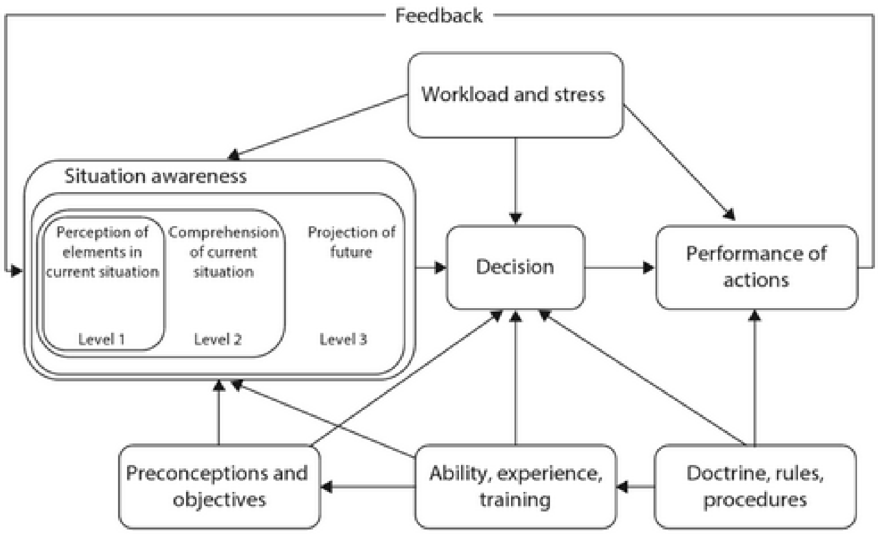
Due to such impactful research, there has been a growing interest in Human Factors training in educational curricula for healthcare, when compared to emphasis and focus towards technical skill acquisition (Lin, Park, Liebert, & Lau, 2015). Cuschieri, Francis, Crosby, and Hanna (2001) surveyed the opinions of 58 master surgeons on the desired attributes and assessment methods of surgical trainees. Within the top attributes chosen were cognitive ability (e.g. decision-making), personality (e.g. emotional stability), and innate dexterity. Although the overarching view was that the impact of technical skill such as dexterity was more important than NTS such as cognitive ability and personality, NTS were highly ranked and useful components on eventual operative competence of the trainees.

3.3 Cognitive NTS- *Situational Awareness*

Situational Awareness (Endsley, 1995) had origins as a concept from military aviation. However, exported to human factor engineering and applied cognition Situational Awareness has been defined as "*knowing what is going on around you*" (Endsley, 2000, p2). There are three complex theoretical constructs on different levels that underpin this term, being Cue Perception (level 1), Perception of Cue meaning (level 2), and the Projection of Cue Status at later temporal points (level 3) (Endsley, 2000; see Figure 2). These three concepts do not have to be linear (Endsley, 2016) whereby issues surrounding situational awareness can occur if there is a lack of understanding of a preceding level, however a plethora of research data correlates with this premise. Understanding complexities of a surgical situation has been suggested to be linked inherently to one's surgical expertise (Durning, Artino, Pangaro, van der Vleuten, & Schuwirth, 2011; Gordon, Darbyshire, & Baker, 2012; Mylopoulos & Regehr, 2011).

Figure 2

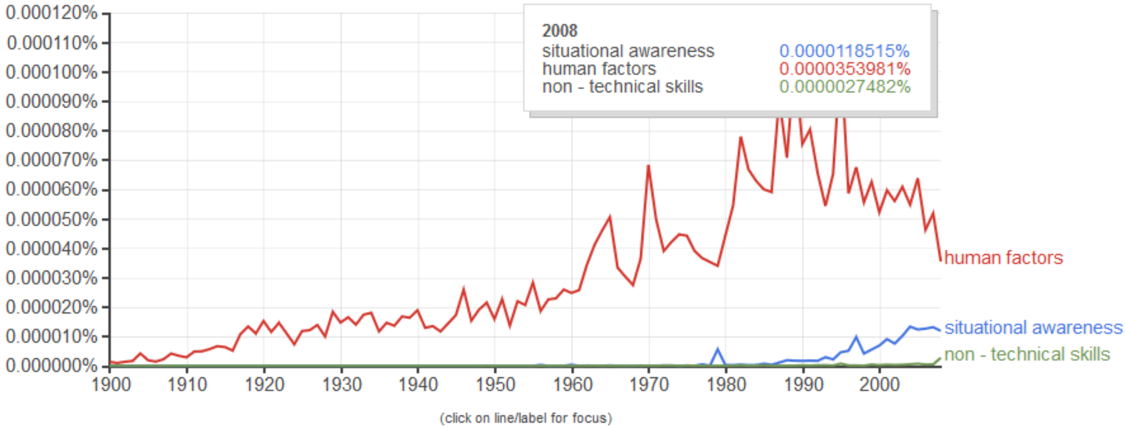
Model of Situational Awareness (From Endsley, 1995). Situational awareness has an information gathering/planning role within the decision-making process.



As an approximation for the increase in research into Situational Awareness, Google Ngram Viewer (books.Google, 2019; see Figure 3) suggested the term 'Situational Awareness' was not significantly included within their book corpora prior to 1988 when compared to 2013. This increase may largely be due to publications by Endsley (1988; 1995; 2000), and Jones and Endsley (1996) which explicitly detailed SA, its importance, and described suggested measures to which practitioners and researcher could better understand and explore the phenomena.

Figure 3

Google Ngram graph (Ngram, 2020) showing the frequency of words used within the book corpora. From early 20th century wartime, Human Factors investigation increased, but it was late 20th century for the cognitive elements of Human performance to be acknowledged.



At the first level of identification in the environment, an individual has basic skills in perceiving the important information within their environment, to form a general and initial mental model of their situation (Johnson-Laird, 1983). The status, features, and dynamic elements related to a task in the environment are perceived to form this basal understanding. Different sets of information can occur for each task, but for every domain perception of information can be collectively sourced through the visual, auditory, olfactory, taste, or tactile senses (Endsley, Bolte, & Jones, 2003). For example, Jones and Endsley (1996) investigated level 1 Situational Awareness amongst pilots and suggested that 76% of Situational Awareness errors were derived from issues in perception of required information from their environment.

At the second level of awareness, a higher level of cognition is introduced. Effortful interpretation of information, combinations of dynamic interactions, and the retention of the transformed information is then compared to the relevance of an individual's goals within their environment. The result is of more understanding of the information in the environment. This can produce a complex ability (compared to level 1) to give more meaning to the situation (through comprehending cues and information). Additionally, it provides more meaning to oneself within the situation through subjective interpretation of the individual's roles, influences, and goals (Endsley, 2000).

At the third level is Projection- being the ability to restructure present information/understanding at a future temporal stage to predict the likelihood of events. This can highlight excellent understanding of the environment and is more frequent in highly experienced individuals with significantly detailed mental models compared to beginners/trainees. Level 3 Situational Awareness has been described as the ability to anticipate future events by cognitively projecting current information (events, cues, dynamics etc.) into the future. Level 3 Situational Awareness does not have a precondition of acquiring both level 1 and level 2 Situational Awareness (Endsley, 1995). However, it may be argued this may occur due to their foundational elements -being perception of the information, to allow the understanding of the collective environment. These fundamentals are ultimately encompassed within level 3 Situational Awareness (Endsley, Boltstat, & Jones, 2003). To achieve higher level Situational Awareness, an individual's mental model requires refining until it is extremely developed and can be accurate in the projection of information.

In ill-defined environments, an individual's highly sophisticated mental model has the capacity to project their current situation into multiple future scenarios to select the most appropriate. For example, a significant proportion of experienced pilots spend more time in anticipating possible future occurrences (Amalberti & Deblon, 1992). This can provide them with improved understanding of the available courses of action, through a longer knowledge gathering period, therefore increased possibility of selecting the most appropriate course of action.However, Cristancho et al (2016) improved upon Endsley's (1995) model as they highlighted that individuals could hold another level of information processing, called Transformation. This entails the technique of considering and refining information that contributes to the creation of new meanings from those pieces of information. Novices interpret information at lower levels, such as for pathology or physiology; however, experts can transform the same information to provide more than one meaning, therefore adding to their repertoire towards making decision. This has suggested that trainees and experts transform the same data differently- yet can be obtainable through training.

Within the environment, changes occur which are linked to a subsequence event occurring. Changes can occur due to an array of inputs, including the human. These changes therefore can be assigned properties to them and each can appropriately be labelled a 'cue' as they represent an indication towards a specific event (Sararit, Haddawy, & Suebnukarn, 2018). The goal of a human is to control or change the environment or situation in a manner that forms a desired event, and it is through the identification of cues towards either desired or undesired events that makes this possible. An undesired event can be prevented when the cues linked to the specific event is not only spotted but addressed in order to stop that cue from reoccurring/developing (Sampene, Littleton, Kanter, & Sutkin, 2019), by rectifying the underlying conditions which facilitated the formation of the cue. Essentially, if humans can rapidly identify cues which have been a prerequisite to negative events, this would reduce the probability of minor errors combining to produce a major undesired event.

If there are issues in an individual's situational awareness abilities, increased risk of a cognitive phenomenon called '*Attentional Blindness'* may occur (Drew, Võ, & Wolfe, 2013; Hughes-Hallett et al, 2015). With significantly reduced situational awareness being a main factor for its cause, Drew, Vo, and Wolfe (2013) investigated 24 expert radiologists into their intentional blindness. The 24 radiologists were experienced specialists in directing their visual attention to detect small abnormalities in different types of images. However, when asked to assess and detect possible nodules in images of lungs, 83% failed to notice a picture of a gorilla, 48 times the size of an average nodule, within the last presented case. Importantly, eye-tracking results revealed they had attended to the spatial location of the gorilla numerous times. However, due to attentional demand of the task and a significant reduction of situational awareness due to their specialist focus, their top-down cognitive processes failed to register the gorilla's presence.

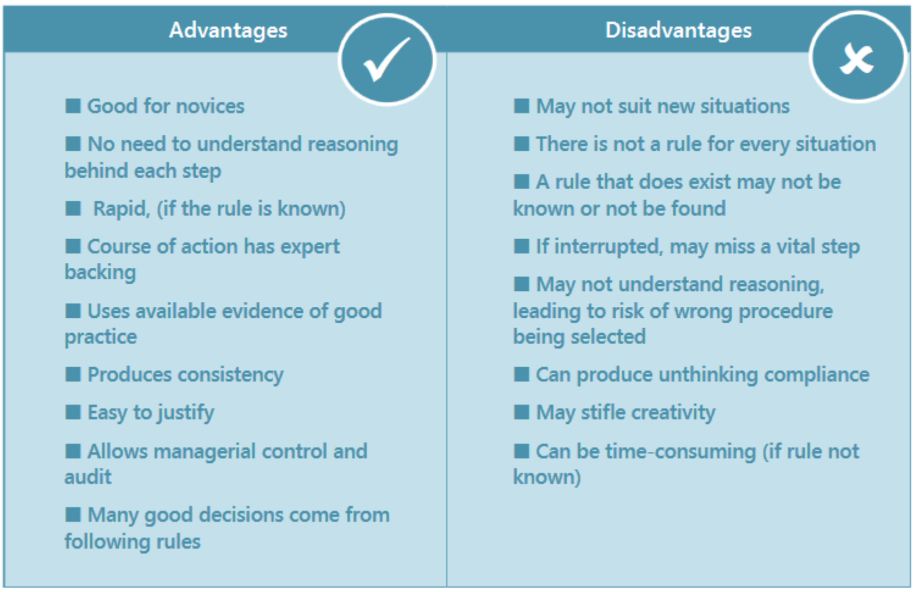
3.4 Cognitive Non-Technical Skill- *Decision Making*

Prominent in surgical research literature (Flin et al, 2007; Sevdalis and McCulloch, 2006) decision-making has been shown to incorporate situational awareness to aid in the final decision-making outcome. It is the ability to recognise the available choices towards different courses of action, to choose the most appropriate option, to implement the decision, and to review the outcome for future improvement. When emphasising the Human Factors perspective in decision-making, further attention is given to how decisions are made in uncertain dynamic, ill-defined, time/resource dependent, and changing ill-structured problems (Gore et al, 2015). Surgical decision making requires accurate real-time response to complex events and understanding of the consequences of each decision. Through compiling the results from previous work (Klein, 1976; Klein & Crandal, 1996; Klein & Zsambok, 1997) and finding a wide display of cognitive processes between individuals, Klein (1989) notably divided decision-making into 4 differing types of underpinning cognitive capacities. These were rule-based, recognition-primed, analytical, and creative decision-making and have been primary components in both early and contemporary decision-making research (Latifi, 2016).

For individuals in their initial stage towards expertise, rules relating to the environment, tasks, and other dynamics are learned initially as rules can be a reliable method to follow to achieve a task. Rule-based decisions are made by the imposition of rules applied by a system that the individual uses to guide their final choice. This method simplifies the situation for the novice however it can require frequent checks of current and past actions to ensure rule are being followed (Flin & Mitchell, 2009; see Figure 4). Nevertheless, this method can allow the individual to fit a situation into the set of rules which describe and handle it best. These rules may be specific manuals or step-by-step guides, and typically written by the experts within the subject area. There are many advantages to rule-based decision-making however as situations become new, more difficult, or form a synergy of different characteristics from different situations then advantages are minimised as rule matching would be difficult (Klein, 2008).

Figure 4

Advantages and Disadvantages of Rule-Based decision-making (from Mitchell, 2013, p 21)



Also known as automatic/automated decision-making, Klein (1989) established the recognition-primed decision-making term through investigating experts in the complex and high-risk decisions of fire commanders (Klein, Calderwood, & Clinton-Cirocco, 1986). Klein (1989) emphasised that these experts were not making deliberate decisions by the previously assumed notion that they were consciously considering each pathway available. Instead Klein et al (1986) extracted their information gathering periods (from video and audio feed) and found the firefighters had much difficulty in explaining their actions. Klein (1989) consequently uncovered that the expert fire commanders used automated tactile, visual, and auditory (and other senses) perceptions to gather information. With reference to prior knowledge that has high similarly to these cues (a prototypical situation), fire commanders decide in a heuristic and compressed/ shortened decisional process and can appear to be automatic due to the speed of response and decision. Intuitive responses usually have time pressures, therefore previously successful schema (relationships of information) of similar situations can be activated, and the relevant action is immediately performed in a '*it worked last time, so let's do it now'* manner (Okoli, Weller, & Watt, 2016). This meant the individual has been in several situations before to have the ability to decide on use of this preloaded response. This stored course of action can also be modified if a situation has minor variations in cues and cue meaning. (Klein, Orasanu, Calderwood, & Zsambok, 1993; Klein & Crandall, 1996).

Analytical and recognition-primed decision-making are the most contradictory to each other (Hammond, 2010). Whilst analytical decision-making takes lots of cognitive effort to identify the situation by analysing all of the environment, intuitive decision-making is automatic and rapid. Whilst analytical decision-making requires one to understand the issues involved before deciding (Crebbin, Beasley, & Watters, 2013), intuitive decision-making requires a small part of the problem and the rest of the environment can be extrapolated by matching similar past experiences (Klein, 2017). Lastly, an analytical decision-making method would generate various courses of action to choose from or by finding help to solve issues at hand as it relies less on pattern recognition but more assessment of each part of the issue (Mishra, Allen, & Pearman, 2015).

The most difficult cognitive abilities to obtain- creative decision-making is formed when experienced individuals have the cognitive abilities to form new options to solve issues that would not be apparent/understood to other professionals in the same situation. Experienced individuals can extract more information from objects in the environment and use this information in innovative ways (Klein, 2017). If no previous consideration was possible, and a suitable response to a situation cannot be accepted by an expert's previous mental models then creativity may be required. Experts would deploy a high cognitive effort to form a new solution (Flin, Youngson, Yule, 2015). Risk is high as the novel method has not been previously tested for rate of success. Time sensitivity may demand a solution in a short timeframe, and individual factors such as fatigue can have greater effects. For example, if a tracheostomy kit breaks or is not present, a surgeon would no longer have the correct equipment to perform a tracheostomy. Yet a surgeon with creative abilities could understand how to modify or recreate the equipment with the tools around them to improvise a similar instrument with the same functionality.

In aviation, decision-based human error attributed to over 70% of fatal crashes (Green, Tsiroyannis, & Brennan, 2016). Similar characteristics involved in surgical decision-making have highlighted the need for effective decisions to minimise errors. These characteristics involve irreversibility, high stakes, and patient outcome uncertainty (Davis, Sevdalis, Jacklin, & Vincent, 2012). Earlier approximations have suggested 25 % of critical events in the OR are related to technical skills with 75% of related to decision making (Spencer, 1978; Evans, 2001). Although present at the pre-, intra- and post- operative stages, the intraoperative stage depends most on the component of situational awareness, which aids effective decision-making which further highlights the dependency on cognitive skills, in addition to technical proficiency.

The use of simulation to teach decision-making has been challenging due to the nature of its non-technical construct. However, through virtual simulation training environments tools can be created to improve some elements of decision-making in surgery. Indeed, there are several recent surgical simulators aimed to teach decisional skills (Flowers & Aggarwal, 2014; Lin, Park, Liebert, & Lau, 2015; Sarker, Rehman, Ladwa, Chang, & Vincent, 2009; Servais et al., 2006). Some of these may have developed from previous acknowledgment that the teaching of decision-making by surgical simulation needed much more consideration to become effective (Anderson, 2012). These simulators have helped to improve the availability of basal decision-making training, yet Vannaprathip et al (2016) stated there still has been insufficient attention supporting the role of perceptual cues in simulator training. Therefore, comprehensive tools that include the role of cue influences upon surgical decisions, should be the next step in developing effective trainers. This is because differences in cue understanding can demonstrate why differences in the types of decision-making can occur.

3.5 Social Non-Technical Skills- *Communication and Teamwork*

Communication has been defined as the transfer of information between 2 or more individuals with the roles as senders and receivers of the information (Wahr et al, 2013). Communication is a facilitative feature for a plethora of other element such as coordination, decision-making, training, cooperation, efficiency etc (Nagpal et al, 2010). Information is transferred by speaking, writing, signing and other behavioural outputs. However, this can be affected by a lack of information, corrupt distribution, or incorrect comprehensions whilst being transferred and received, these are the causes of communication error.

From simply handing over an instrument to recovering from complex unpredicted events, these goals require transfer of information before, during, and after actions are performed. The success of completing a goal depends upon the information transferred. Too little information or lack of quality, and the meaning of the goal may not be understood. Inaccurate or ambiguous information may distort understanding of the goal and the outcome may be different to what was expected. In the domain of aviation, the standardization of communication has occurred whereby a set process is used to instruct how to transfer information, called the Crew Resource Management (Salas, Burke, Bowers, & Wilson, 2001). This allowed information to be transferred in the optimal format to keep the desired information precise towards fulfilling the goal. Comparably, in more simple information exchanges in surgery, a consistent method is typically applied. For example, a surgeon asking for a scalpel would identify the receiving individual, identify the tool of interest, and transfer instructions of who/how to pass it. However, for complex goals surgical information transfer has higher risk of miscommunication, and an explicit standardised method has not been fully integrated to precisely relay packages of information that are required to successfully complete a goal (Raheem et al, 2018).

Measures in communication skills suggested they have been the weakest OT behaviour (Wahr et al, 2013). Morris et al (2003) investigated the underpinning surgical system failures which consequently resulted in insurance payment due to error. They posited 87% of failures were due to issues in communication. However, the Joint Commission report (2012) also stated communication failures were root cause for 65% of unanticipated events (see also Mushtaq et al 2018), additionally communication failure was a lead contributor to teamwork failures as a similar figure of 89% of teamwork issues stemmed from communication complications (ElBardissi, Wiegmann, Henrickson, & Wadhera, 2008). With a general summary of communication Gawande, Zinner, Studdert, and Brennan (2003) also suggested incomplete or erroneous communication was a causal factor in 43% of errors made during surgery.

Cues and signals released by distractors within the environment can affect communication. The OT possesses a surgical workflow which may be disturbed by the presence of one or more distractors/interruptions (Weber, Catchpole, Becker, Schlenker, & Weigl, 2018). These may be auditory such as phones, pagers/beepers, conversations both within and between OT attendants either with task relevance or no task relevance. They may be from visual cues such as unnecessary movement, people walking in/out, equipment transfers etc. These can contribute to flow disruptions such as equipment and technology issues, teamwork/communication failures, and training-related disruptions (El Bardissi, Wiegmann, Dearani, Daly, & Sundt, 2008). However, many task-based and technical surgical simulators have not incorporated training associated with higher level cognition to manage the influences of distractions on planning, decision-making, and problem solving (Shiralkar, 2011). The identification, handling, and impact reduction of distractions upon an individual's performance and cognition has had little simulated training materials created, that can be easily accessed to a wide audience therefore more practical training materials are required.

Indeed, modern methods of surgery have increased need for communication training. Raheem et al (2018) were the first to investigate and analyse communication during robot-assisted surgery. Their study amplified the importance of communication due to the unique structure of robot-assisted surgery as the surgeon placed away from the team, and with more equipment. Interpersonal cues may be lost and increased dependency on verbal information increases the need for effective information. Resultant loss of contextual or specific directional information can increase a result of ambiguity.

In the example given of a surgeon asking for a scalpel, the individual initializing communication should explicitly identify who they are addressing. If an individual has not been addressed there may be mix up of task allocation. For example, '*can someone get me another drill please*?' does not direct attention to one individual and distributes task allocation. Many resultant issues can occur because of such non-directional phrasing. The common result being an assumption that someone else will execute the task, ironically leading to nobody performing the task. Other results may be that more than one person will complete the task, additional time spent finalizing who will complete the task, and a closed-loop communication system may not be used. A closed-loop system being that the receiver can repeat and confirms they are singly assigned and understand the task, usually executed 3.6 times faster than open loop communication (Robson, 2017).

To address this, the 'Theatre Cap Challenge' (Psnetwork.org, 2019) has been one attempt to improve use of name and role in communication transfer. Using an individual's name transfers accountability of the information, however OT staff may not know, or forget colleagues' names. By adding visible identification onto their caps in the OT, name and role introduction in theatre has increased from 38% to 90% (Burton, Guerrerio, Turner, & Hackett, 2018; El-Shafy et al, 2018) and closed-loop communication was more likely. However, the uptake of this practice has been slow due to usability, security, and hygiene hurdles. Instead, orientation towards the modification of behaviour to support closed-loop and unambiguous communication could be easily incorporated into future simulations and may improve the issue with less resource usage. Again, this could be comparable to other domains such as firefighting (Gabor, 2015), or aviation who use structured language to ensure the necessary information is included in information transfer (Holtaway & Jackson, 2017).

Christian et al (2006) investigated patient safety in the OT when referring to communication efficiency and information flow. They observed that transfer and flow of information has vulnerable points to which patient safety can be compromised. From this, Halverson et al (2011) implemented a communications training module to OT personnel that demonstrated information transfer techniques. Based on the Crew Resource Management (Gross et al, 2019), some techniques were to call out critical information, and to repeat back any verbal requests in a closed-loop and non-ambiguous method, along with a standardised approach to the 'hand-off' event when responsibility for the patient's care is transferred to another individual/team. They summarised that such a program to teach communication skills in the OT can improve information transfer by changing the strategies in use, as there was a significant reduction in communication error frequency.

Teamwork has been described as the combined behaviour from two or more individuals engaged in the required activities necessary to compete one or more tasks (Gardner, Kosemund, Hogg, Heymann, & Martinez, 2016). Teamwork can be viewed as the composition of the individual's ability to utilise their communication, leadership, and situational awareness to convey information or produce actions which support the team goals. One substantial ability that strengthens teamwork dynamic is Anticipation. Lack of situational awareness is one of the primary events which account for the occurrence of human error in anticipation a high workload setting (Kaber & Endsley, 2004). Teamwork is a vital element in avoiding this occurrence by helping to share situational awareness with colleagues who may have temporary reduced awareness. For example, in micro-neurosurgery most of the time the surgeon will allocate attention to the microscope and have less recognition of everything outside of this. Additionally, the microscope can act as a barrier of communication between colleagues and extra effort is required to provide a shared understanding of the previous, current, and next steps towards their goal (Afkari, Bednarik, Makela, & Eivazi, 2016). If teamwork degrades there is risk that an individual may not get the necessary information from their team to support their awareness, therefore decision (which relies on SA) may be made which would increase chance of events which move away from the team goals.

3.6 Social Non-Technical Skills- *Leadership*

Leadership has been defined as "the process of influencing people towards achievement of organizational goals" (Naylor, 2004, p.354). Flin, Yule, and Youngson (2015) compacted leadership to the 3 point of setting and maintaining standards, coping with pressure, and supporting others. To be more quantifiable Parker, Flin, McKinley, and Yule (2013) held ten focus groups to discuss and extract leadership behaviors to form a taxonomy. They identified 8 elements: maintaining standards, managing resources, making decisions, directing, training, supporting others, communicating, and coping with pressure. This formed a general description of leadership both in and out of the healthcare domain.

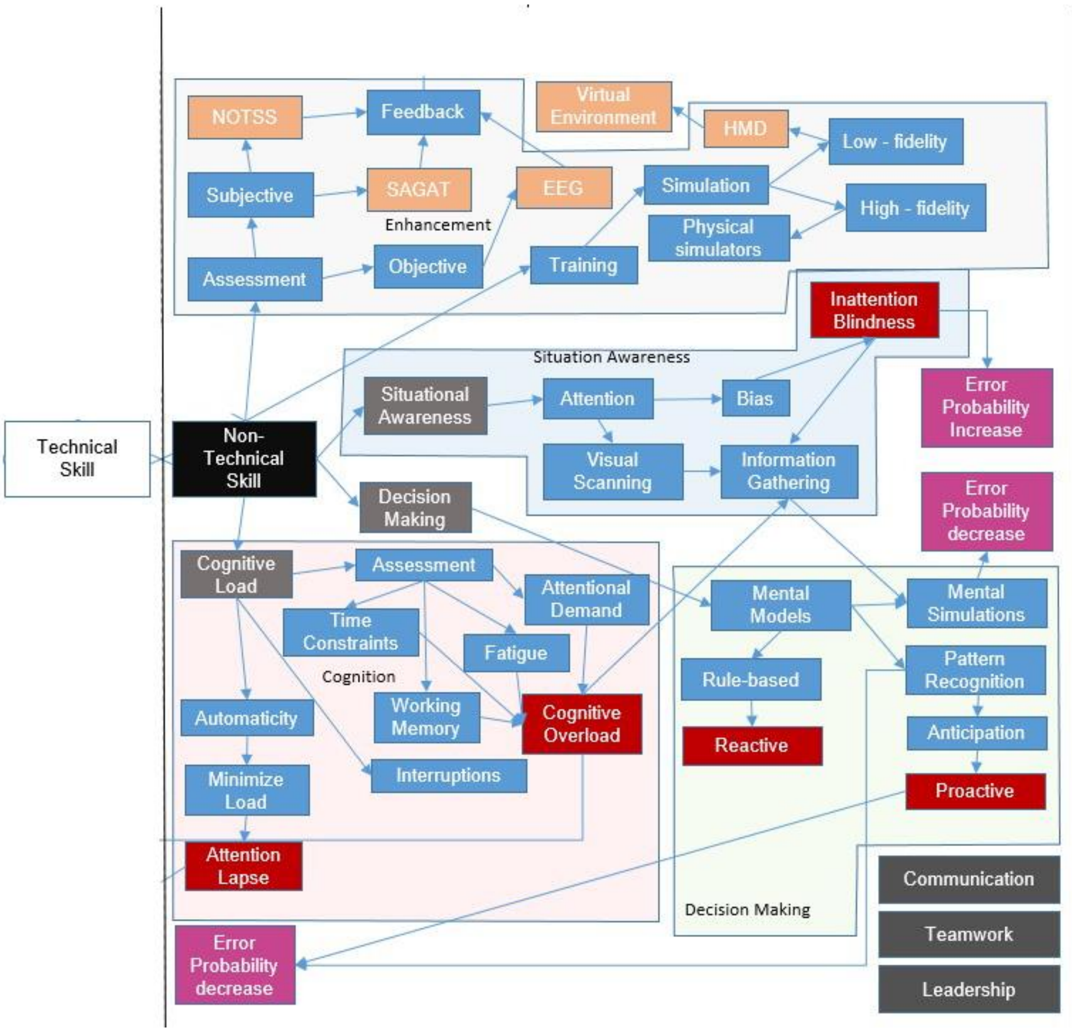
Leadership has been used as a non-technical strategy to prevent the interruption of communication, co-ordinated care breakdown, enhance safety behaviours, and can be a drive towards goal achievement and task completion (Fletcher et al, 2002; Flin et al, 2003). For medical students' overall healthcare delivery there has been a decrease in medical error rates, improved communication skills, and improved patient trust/satisfaction (McKnight, et al, 2002; Awad et al, 2005). Leadership has many trait, behaviour, or power-based models and theories depending on the domain and contexts (I.e. Fiedler, 1967). However, an encompassing theory for healthcare individuals that has had most success was the model formed from the combination of both transformational and transactional theories of leadership (Bass, 1990; also see Jamison, 2019). This model has facilitated research to recognize the personal characteristics that are required for optimal objective outcomes and provide measures of assessment when paired with the 8 elements suggested by Parker, Flin, McKinley, and Yule (2013). For example, transformational leadership can be identified by the intellectual stimulation of others towards their collective goal and their ability to consider the needs of the individuals as well as a group to optimize productivity. Reward and praise are in the form of encouragement to develop the strengths in their team, and promote independent motivations so everyone can fulfill an effective team culture and environment.

However, when understanding the transformational/transactional theory of leadership (Jamison, 2019), the effects of transactional behaviour when in a leadership role can be limited. Being task-focused it delegates responsibility for tasks with reward for target achievement and scrutiny when mistakes or failures occur. Hence, this leadership style is a ‘back and forth’ style of conditional positive attention if tasks are successful. This can be problematic when tasks may not be fulfilled as planned. For example, a team can be susceptible to exhaustion in order to achieve the desired tasks (Perry, Witt, Penney, & Atwater, 2010). Siu, Maran, and Paterson-Brown (2016) suggested that poor leadership skills were strongly correlated with preventable intraoperative incidents, based on observational analysis. Hence with more complex systems and dynamic interactions in surgical and medical settings there may be greater risk of suboptimal leadership strategies.

The roles of non-technical and applied cognitive skills has been identified as a significant factor in the patient safety system. This insight was gained from reviewing the state of current technical training along with non-technical skills research and brief theoretical outline. Up to the current chapter, a simplified outline of the researcher’s early understanding can be perceived in Figure 5. Based on this, a review of non-technical skill training within surgical curricula was required to ascertain limitations in its existing implementation and identify the most challenging areas.

Figure 5

A simplified approximate landscape of the researcher's early understanding of the Non-Technical elements and their dynamic links.



4.0 Contemporary Training Methods of Situational Awareness

Previous training interventions into situational awareness have had general success. Moorthy et al (2006) provided 20 surgical trainees with high fidelity surgical team crisis teaching that included assessment pertaining to the NTS between junior and senior trainees. Their results suggested NTS skills including situational awareness were not significantly different between the two participant groups after training. This suggested simulation which included specialised feedback can be successfully deployed when a practical method of simulation and resources of expert involvement are used. With similar methodology to Moorthy et al (2006), Menhadji et al (2013) provided a Minimally Invasive Surgery (MIS) team of 16 with a high-fidelity simulator for surgical crisis training. NTS of junior and senior urology residents were compared during simulation. Differing to the results from Moorthy et al (2006), Menhadji et al (2013) identified significantly lower individual situational awareness scores among junior residents compared to seniors. These scores surrounding information gathering, projecting and anticipating future events, coping with pressure, and selecting/communicating options. They concluded appropriate construct validity of the intervention for situational awareness training at an individual level. Additionally, all urology residents stating its usefulness in developing communication to make residents aware of unlikely but potentially critical OT errors, development of simulation was suggested.

Flin et al (2007) also expanded on the foundation from Moorthy et al (2006) to provide a short NTS course for all skills, not only situational awareness. The workshop consisted of videos and lectures relating to all aspects of NTS, to 21 surgeons. The summative feedback of the course entailed the surgeons' evaluation based on their attitudes, content relevance, and usefulness. They concluded that the course had good face validity and the majority believed it was beneficial as it encouraged behavioural change in terms of NTS in the OT at the individual level. With growing interest and reiteration of this material, teaching hospitals started to incorporate modules into their Intercollegiate Surgical Curriculum. For example, the Oral and Maxillofacial Surgery programme (ISCP, 2015) syllabus provided an indication of the major components of the core surgical training programme. The criteria to be accepted for intermediate and final stage training modules included 'Personal Skills'. The skills were broken down into Judgement under Pressure, Communication Skills, Problem Solving, Situational Awareness, Decision-Making, Leadership and Teamwork, Organisation and Planning.

4.1 Contemporary Methods of Training for all Non-Technical Skills

However, after 2015 there has been scattered deployed and varying Human Factors training, of which includes the training of NTS. A plethora of organisations have attempted to create their own materials from the same foundation (Terema, 2020; healthcareconferenceuk, 2020). The World Health Organisation (WHO) encouraged patient safety courses into teaching curricula, it has a module detailing the human role in healthcare for students to understand their effects of behaviours in patient safety (WHO, 2019). The NHS Health Education program (westmidlandsdeanery, 2020) have mandatory Human Factors training for core surgical trainees.

However, being the most successful in its deployment and empirical results towards patient safety, there was an increase in use of the *Non-Technical Skills for Surgeons* framework (see Flin *et al*, 2015). Recently, the predominant training available was from the Royal College of Surgeons Edinburgh (rcsed.ac.uk, 2020) and oriented around using the *Non-Technical Skills for Surgeons* framework (NOTSS; see Flin, Youngson, & Yule, 2015). The Royal College of Surgeons of Edinburgh (RCSE) Patient Safety Board has educated senior trainees and consultant surgeons, with junior surgical trainees being able to access e-learning NOTSS modules. This framework has close ties to NHS England, and it has become mandatory in many institutions that have implemented it.

The Royal College of Surgeons developed a two-day course which accommodated for most healthcare roles. Constructed around the work by Flin et al (2006) and Yule et al (2008) this was one of the first, and the most comprehensive training session available. To understand the most up-to-date real-world implementation, the researcher of this project attended 2 Non-Technical Skills workshops in 2016 and 2018, being invited to be an observer. Located in Edinburgh in 2016, the first course was attended by surgeons and anaesthetist from an array of differing departments. After the session, the researcher probed the individual leading the workshop, to understand the rationale for approaching non-technical skills training in their specific training format. Due to the hierarchical structure of healthcare institutions, the implementation of new materials into an institution would have to flow from a top-down approach. Consequently, high-level individuals can change surgical training methods within their institutions.

In 2018 the course had been established in the University of Leeds and the content had not significantly changed since 2016 as underlying theory and human factors content was presented (Medicaleducationleeds, 2020). The Royal College of Surgeons Edinburgh have performed workshops around the world, but largely in areas where their impact would be most effective. For example, locations in India such as Hyderabad and Dervan, have held the NOTSS workshops (rcsed.ac.uk, 2020b). Other regions worldwide have also utilised the NOTSS material and non-technical workshops. The Royal Australasian College of Surgeons (surgeons, 2019) has provided monthly NOTSS courses and other regions such as Japan (Tsuburaya et al, 2016), have incorporated them into their respective institutions with success. This rising implementation has highlighted the growing awareness of the importance of NTS and acknowledgement of the benefits of incorporating NTS training into syllabi and curricula. Certainly, future NTS *formative assessment* has consideration from the Intercollegiate Surgical Curriculum Programme (ISCP, 2015). This has suggested that there is potential for forthcoming NTS training enrolled for differing trainee skill levels that may involve benchmarks to ensure the appropriate skill levels to correspond with their overall skill progression.

4.2 Functional Impact for Non-Technical Training

With a range from novice to experts, revisions of functioning surgical systems and appropriate procedures have become an emerging practice to assess, identify, and improve the quality of non-technical methods. Implementation of routine video recordings for select surgical procedures has captured performance in order to reduce or eradicate low standards of code and conduct (Birkmeyer et al, 2013). Additionally, Lavoie, Cossette, and Pepin (2015) assessed 15 critical-care experts and 234 bachelor level critical-care nursing students by using a high-fidelity simulator under different conditions. They were able to illustrate that simulators are useful for typical training, but also to assess individuals' perceptive, cognitive, and decision-making processes. Lavoie, Cossette, and Pepin (2015) concluded that students perceived the simulator experiment as a tool to help them realise what they forgot to assess or did not notice, in their initial task assessment of their simulated environment. This in turn improved situational awareness and benefitted in better decision-making, judgment accuracy, and therefore decreasing chance of error.

Similarly, as there has been recognition of restricted opportunities to train in oral and maxillofacial surgery (Goodson et al., 2013), Servais et al (2006) investigated the impact of a specifically designed Virtual Learning Environment pertaining to common maxillofacial emergencies. Their rationale stemmed from the indication that although there has been considerable literature describing the benefits of simulation in general surgery, there has been scarce information specifically in its use in oral and maxillofacial surgery residency training (Coffey-Zern, Calvi, Vorrasi, & Meara, 2015). They provided a free Virtual Learning Environment for one month with pre- and post- training assessments, which measured confidence and knowledge as subjectively determined by the 29 junior A & E doctors who participated. Both confidence and knowledge of maxillofacial surgeries showed significant improvements (+28%, $p<0.001$). This highlighted that postgraduate medical education can successfully incorporate Virtual Learning Environments both in standalone conditions and as part of a course (Cook, Erwin, & Triola, 2010). These emerging practises have highlighted a rising need for efficient training methods and effective non-technical (with inclusion of technical) assessment framework for the future surgeon.

4.3 Disproportion of Non-Technical Simulators - Influences

Applying Human Factors research has been able to provide partial explanation for the difficulty in NTS inclusion in Surgical Simulators. For experts, when a situation requires their attention, their information gathering and skillset combine information different to trainees due to their experiences (Vannaprathip et al., 2016). Indeed, although expert surgeons are aware of different decisions, generally they are aware of around 70% of the *reasons behind* decisions during crucial procedures (Pugh, Santacaterina, Da Rosa, & Clark, 2010). This has been suggested to be due to significant improvement in cognitive abilities over a given time period creating 'automated cognition' (Sullivan, Yates, Inaba, Lam, & Clark, 2014). Evidence for this phenomenon details that more preventable surgical errors have occurred during routine complex surgery when compared to emergency or less common procedures (Regenbogen et al, 2007). As expert surgeons rely on their automatic decision–making responses when performing surgery, they are inefficient at deconstructing their cognitive processes into steps (Moulton, Regehr, Lingars, Merritt, & MacRae, 2010) as significant compression of the decision-making process has occurred.

Yet, the Halstedian apprentice model has been the tradition in surgical skills training (Palter & Grantcharov, 2010). Whereby surgical experts provide close supervision to trainees when performing a task as they can convey how they interpret a situation (i.e., filtering unnecessary information), and how the trainee can later implement these skills. As experienced surgeons have cognitive abilities that are significantly different to trainees, Human Factors research can help bridge this gap in training. It enables expert cognition to be extracted, decompressed, and implemented into a delivery method most suited to the needs of modern trainees. With the lack of cognition-focussed Human Factors incorporation in surgical training, there can be difficulty in completing this process. This may be a reason why NTS simulators are disproportionate to simulators of other skills.

4.4 Shortage of Non-Technical Skills Training

Although there has been strong support for non- technical skills improvement, Graafland, Schraagen, Boermeester, Bemelan, and Schijven (2015) evaluated nine journal articles that assessed surgical team training in simulated environments. For this, they reviewed literature pertaining to the validity of methods for improving non-technical skills in the surgical theatre. They conversely highlighted that more research was required to incorporate cognitive factors in surgical teaching curricula, and that errors can still be due to non-technical factors, such as incomplete understanding of the situation (Fabri & Zayas-Castro, 2008). Importantly, this can be independent of a surgeon’s technical skills (Tang, Hanna, & Cushcieri, 2005; Dankelman & Grimbergen, 2005). Graafland, Schraagen, Boermeester, Bemelan, and Schijven (2015) suggested non-technical skills such as situational awareness needs more attention in terms of training prior to the OT to improve patient safety. Consequently, they suggested a necessity to investigate how non-technical skills differ between individuals comprising of various technical abilities within their field, and how this translates to influence learning capacity, communication, and system interaction.

Brunckhorst et al (2015) investigated existing urological training practices in order to determine the role of different types of simulation, including virtual simulation. They similarly concluded an existing under-utilisation of simulation, notably virtual simulation, for trainees. Indeed, both trainees and specialists established technical and non-technical skill simulation had a role within current training curricula, as it would increase patient safety. In agreement, Piromchai, Avery, Laopaiboon, Kennedy, and O’Leary (2015) pointed out “*there is a paucity of studies examining non-technical skills*” (p26) and suggested a possible reason for this may be due to the types of virtual simulations that have been available. Simulations have targeted psychomotor and procedural skills, resulting in limitations in the ability to assess non-technical skills and team interaction.

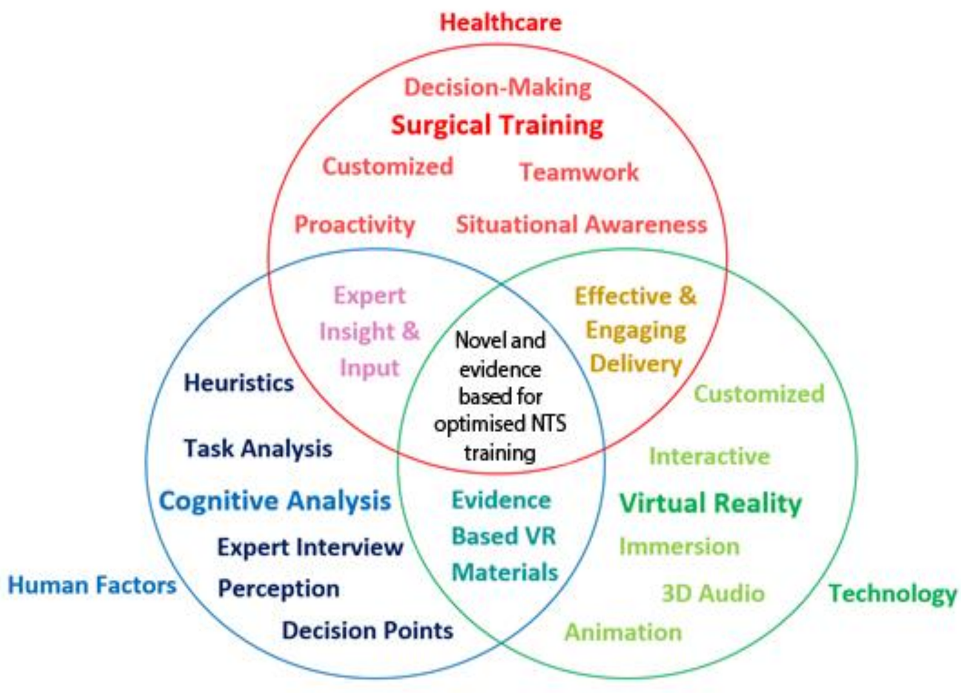
4.5 Future of Virtual Reality Surgical Training Devices

From the previously discussed research, this project identified a need for the merger of advanced VR simulation and NTS training. An excellent reason for rejection of this idea is from Yiasemidou et al (2016). They suggested that although 86.67% of surgeons agreed that simulation could be used for non-technical skills acquirement, the remaining 13.33% were either indifferent or disagreed/deemed not important. Additionally, 33.33% disagreed or deemed it not important that simulation training prior to performing a procedure was necessary, with 13.33% being indifferent. This meant homogenous support was not found for the role of virtual reality simulations. To justify the position of the research that there is a need for the merger of advanced VR simulation and domain specific NTS training, a critical review of Yiasemidou et al (2016) found a methodological limitation. This was the absence of a new type of advanced virtual surgical simulation in their evaluation of simulation suitability. This has also been the case in all of the relevant studies previously mentioned in this literature review up until this point (except Nakayama et al, 2016). Although a high fidelity computerised virtual simulator was compared by Yiasemidou et al (2016) to three other simulations, the recent virtual reality tools has yet to be assessed in terms of its efficacy for the innovative synergy of technical and non-technical surgical skill acquisition.

Therefore, the following chapters in this thesis aim to explore the role of expertise in NTS training. The methodologies used to extract knowledge and form them into training materials was investigated in Chapter 5 and focused on the *Human Factors* circle within Figure 6.

Figure 6

Venn diagram identifying the mergers of separate themes. The combinations of these three themes may be a framework to promote pedagogical materials for accelerated education of non-technical skills for surgeons.



Chapter 5 - Human Factors Methodology and Expertise for Non-Technical Training

5.0 Advantage of Usage for Surgical Training

Human Factors integration can support surgical experts to articulate tacit operative steps and cognitive decisions. Experts perceive and reprocess their experiences to improve upon meta-cognitive abilities and mental representations. Human Factors research can extract, decompress, and repurpose this knowledge into a delivery method suited to modern trainees (Hoffman, Crandall, & Shadbolt, 1998). Newly created training methods can be successful when identification of experts' cognitive processes are included (Schaafstal, Schraagen, & van Belro, 2000; Velmahos et al, 2004).

5.1 Experts in the Human Factors Process

In most complex domains, there are individuals who have differing skill levels. The spectrum varies from those who are learning things for the first time, to those who have been in the domain for several decades (Hutton & Klein, 1999). Navy admirals, historical motorsports champions, chief firefighters, renowned surgeons, or large-scale business magnates likely have had years (usually decades) of deliberate, in-depth, and complex experiences. Those new to a domain may have to endure fails, issues, misunderstandings, and losses but also will achieve goals, solve problems, and hone their skills to eventually reach a similar level of expertise. Achieving these cognitive skills faster may help not just the individual, but their group and ultimately improve their environment (Suss & Ward, 2018). Therefore, these domain-specific specialists hold skills that allow for significantly improved performance compared to non-experts. They have superior monitoring abilities for both their own motor and cognitive actions, and the actions of others (Ericsson & Lehmann, 1996). For example, when tested in surgical simulations the high level of outcome performed by experts may be used as a benchmark performance. Rules and guidelines previously used are refined overtime that experts will not rely on them, and possibly deviate outside of these boundaries due to stronger intuition of each situation that has better relevance than the protocol they are typically meant to follow (Eraut, 2000).

5.2 Defining a 'Subject Matter Expert'

Dreyfus and Dreyfus (1996) proposed that individuals can be categorised as they progress through the Apprenticeship Model. This model described an initial '*novice*' which is succeeded by an '*advanced beginner*', progressing to gain '*competence*' – these are surpassed when an individual is deemed '*proficient*' along the skill spectrum. To be distinguished in the higher-level categories of the skills a '*Subject Matter Expert'* (SME) and '*Master*' will have obtained exceptional knowledge to the antecedent categories, such as contextual flexibility (Selxas-Mikelus et al, 2014; Zhou et al, 2018) along with improved procedural and declarative skill. Sociometric analysis (aka social criteria), extended domain experience, performance analysis, and education/career analysis have been the typical classification criteria for individuals when categorising expertise. The requirement of the indicative 10,000 hours of practice of surgery to form expert proficiency (Schaverien, 2010) has popularity in research (Ericsson, Prietula, & Cokely, 2007; Ericsson, Krampe & Tesch-Romer, 1993; Hirschl, 2015). However, the relationship between clinical outcome and the amount of operation practice does not determine skill level. Similar proficiency can occur in individuals with a variety of operative experience. It has been proposed that significant increase in knowledge and experience have a more direct role in understanding more than the amount of time performing surgery.

Due to such inconsistencies with the classification of an expert within a system, researchers have frequent issue with recruitment and the quality of the insights gained by nominated SME's (Klein, Borders, Newsome, Militello, & Klein, 2018). As a human factor researcher communicates with clients, the meaning and terminology of an 'expert' may be interpreted differently by both parties. An expert from a cognitive perspective may be an individual with richer mental models, increased tacit knowledge, and improved abilities to anticipate future states (Suss & Ward, 2018). Yet (although with typically positive correlation), an individual's rank, years of service, or leadership style can be misinterpreted as the criteria by the domain client to then erroneously select an 'expert' which may not meet the researcher's needs.

5.3 Cognitively Advanced Practitioners

Therefore, a change of terminology was made in the current project to alleviate these discrepancies. A new term was created for this project, whereby a Cognitively Advanced Practitioner (CAP) was referred to as the 'expert' required in the context of Human Factors and applied cognition. 'Expert' and 'CAP' were used interchangeably based on this definition. For example, after months of training an apprentice air traffic controller may now comprehend what the icons and symbols mean on their screen and how these represent the pilots, passengers, and airports along with how they communicate to each other via a specialist system (Wickens, 2019). The controller may fall on the spectrum between *'beginner'* and '*competent'* at this point. Over time, the comprehension of a system becomes so strong that they can manipulate the same entities in unique ways that may appear irrational to others, perhaps to achieve a goal faster, avoid an uncommon error, reveal embedded patterns, or anticipate future situations. Achieving this proficiency, the controller may be regarded as a Cognitively Advanced Practitioner (CAP).

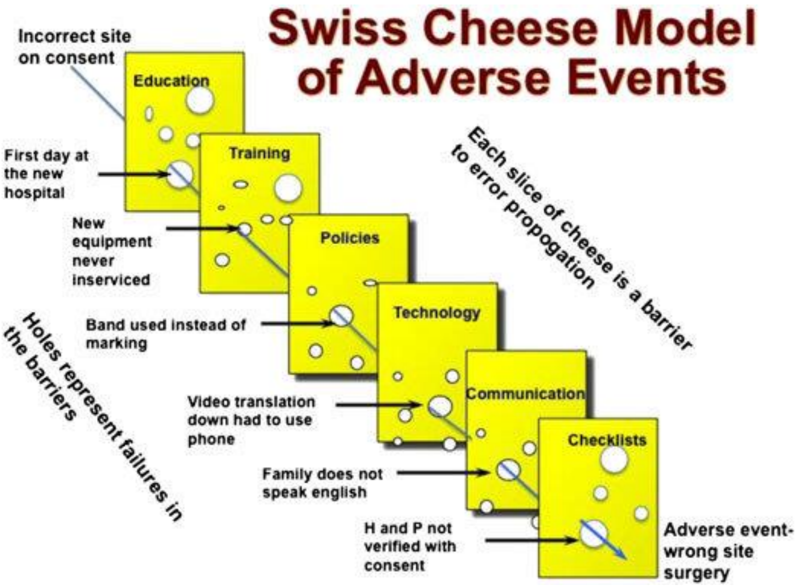
5.4 Relevant Theory and Methodology for Healthcare

5.4.1 Swiss Cheese Model

The utilisation of CAP's in Human Factors research has importance in healthcare to fulfil the overarching goal of improving patient safety and outcome. Expert insight is valuable to strengthen the barriers which avoid undesired events occurring. This overarching goal can be explained by the Swiss Cheese Model (Reason, 1990). The model has been primarily used to understand and explain accident causation and can highlight the dependency on expertise to help customize its components specific to each institution. The premise of the model has been that flaws exist in the many layers of defence/prevention in a system, which may add up to major error(s) if all defences occur simultaneously (see Figure 7). The analogy with reference to Swiss cheese represents the human/environment system comprising of multiple layers. This analogy has an underpinning meaning in that under specific conditions, the holes found in slices of Swiss cheese can be aligned to compile an array of holes creating a gap with no resistance/material. An item would be allowed to pass through this formed channel. Each layer of cheese represents one individual safety net or barrier that is implemented into a system, holes in each slice represent the opportunity for a minor error to occur due to a limitation in the safety setup. These errors can combine to form major undesired events due to the combination of 'trajectory of accidents' (Reason, 1995). However, within the dynamics of real-time complex systems, it is possible to change this trajectory which may decrease the real-time probability of the alignment of the subsequent layers. To further implement human factors into healthcare, the Swiss Cheese Model has been transferred from aviation (Wiegmann & Shappell, 2017) and adapted for safer surgical settings (Flin & Mitchell, 2008; also see Gerstle, 2018).

Figure 7

An example of the Swiss Cheese Model being adapted to provide an example of a surgical resident's wrong site surgery (From Seshia, Young, Makhinson, & Smith, 2017).



5.5 Human Factors Analysis and Classification System

An expansion of the Swiss Cheese Model named the Human Factors Analysis and Classification System (Chauvin et al, 2013) (HFACS) stated there can be four areas of compromise within a system. These are organisational influences, unsafe supervision, preconditions to unsafe acts, and unsafe acts themselves. The HFACS asserted that one or more occurrences from four failure domains are the causation of an incident. Indeed, Green, Tsiroyannis, and Brennan (2016) applied the HFACS relevant to practicing surgeons and suggested that organisational influences such as process and resource management, communication, training and recognition of human factors were possibly responsible for error. Unsafe supervision may be from loss of situation awareness, especially if not recognised by a clinical team, inadequate supervision, or failure of the team to react appropriately if change occurs. Preconditions to unsafe acts were suggested to be from environmental factors, distractions such as noise, fatigue, nutrition of staff, emotional influences, tiredness or boredom, and problems in communication.

These events include both active and latent failures whereby an active failure is the specific error, for example miscommunication in surgery. Latent failure has been defined as resultant from process or organisational error.

For example, latent causes of errors may be increases in work pressure and longer hours. If an error occurs in aviation the pilot as well as the passengers will be affected by the experience. Surgical staff appear to differ from the pilots and other personnel as they would not usually be harmed, however they can sustain significant psychological effects that may devastate both the individual and the teams they are involved within (Sugden, Athanasiou, & Darzi, 2012).

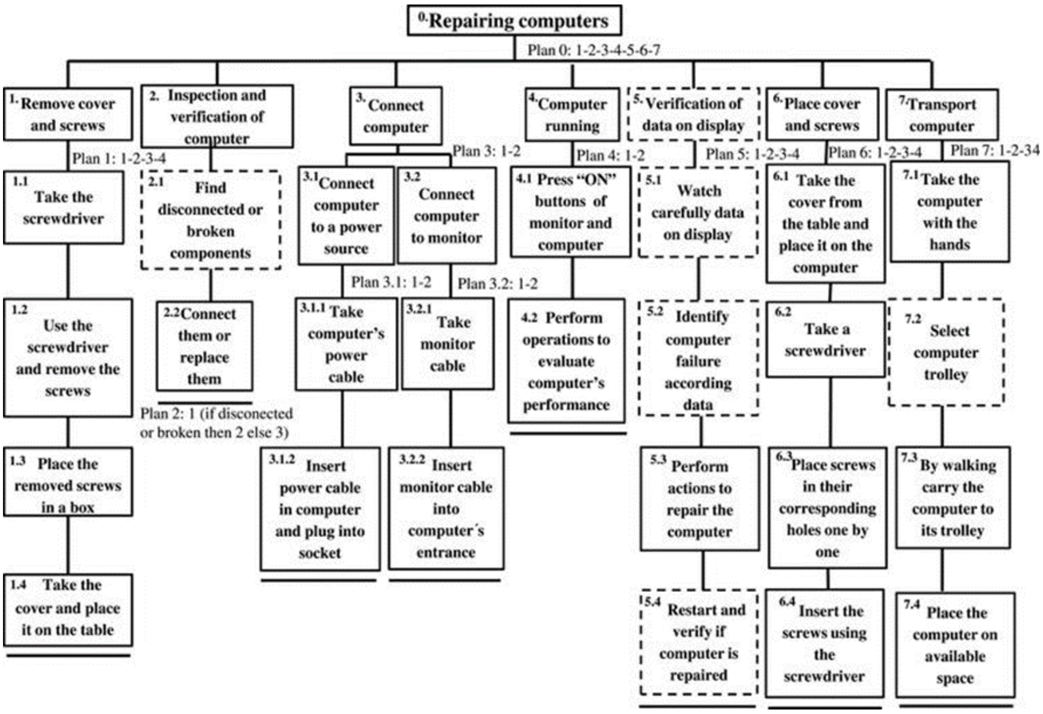
5.6 Human Factors Methodology

To extract valuable feedback from such individuals who have compressed their cognition within their domain, a method of system analysis and cognitive probing is required. As researchers have investigated a wide array of domains, each with specific challenges, personnel, and end goals a plethora of methodologies have been developed to collect data. For example, the SESAR (Single European Sky ATM Research) project have compiled 50+ tools and techniques which orient around measurement of human and system performance (Eurocontrol, 2020). Of which a large majority can be adapted and implemented in other fields. To initially understand a system, one of the most popular methods applied has been a Hierarchical Task Analysis (HTA; Shepherd & Stammers, 2005).

For a researcher unfamiliar with a system, a starting point to firstly understand the overall events and processes within a system can be completed by appraisal of the overall tasks through Hierarchical Task Analysis. It can be used to provide a comprehensive description of tasks to be a foundation to further analysis. A HTA output can be tree or tabular in diagrammatic representation (see Figure 8) and the overall goal is positioned at the top, being the final desired outcome. From this, meaningful sub-goals are formed which, once completed, combine to achieve the overall goal. These sub-goals are once more broken down into further sub-goals and operations (physical or cognitive actions required) until the bottom level HTA branch is reached. This branch level consists only of operations as they specifically detail what actions needs to be performed. Everything above this level specifies goals (Stanton et al, 2013). Lastly, cognitive results produced at later stages of participant interaction can be linked back to the task processes.

Figure 8

An example of a basic Hierarchical Task Analysis of the main tasks in repairing computers. (From Maldonado-Macias, Realyvasquez, Hernandez, & Garcia-Alcaraz, 2015).



Another example of one focused method of knowledge elicitation is a multi-stage interview technique named Concepts, Processes, and Principles (CPP; Clark et al, 2010). CPP captures both unconscious and automated knowledge from CAP's by asking them to list performance sequences in all of the sub-tasks identified in the knowledge representation stage. As a result, an in-depth sequence of a task is created allowing the researcher to then probe into the events of each task. Task questions may consist of the concepts, processes, and principles that an individual may use as a conceptual basis for their approach to the sub-task, or the decision made in the same sub-task. Additionally, the criteria necessary to accept one decision and reject the others can be found.

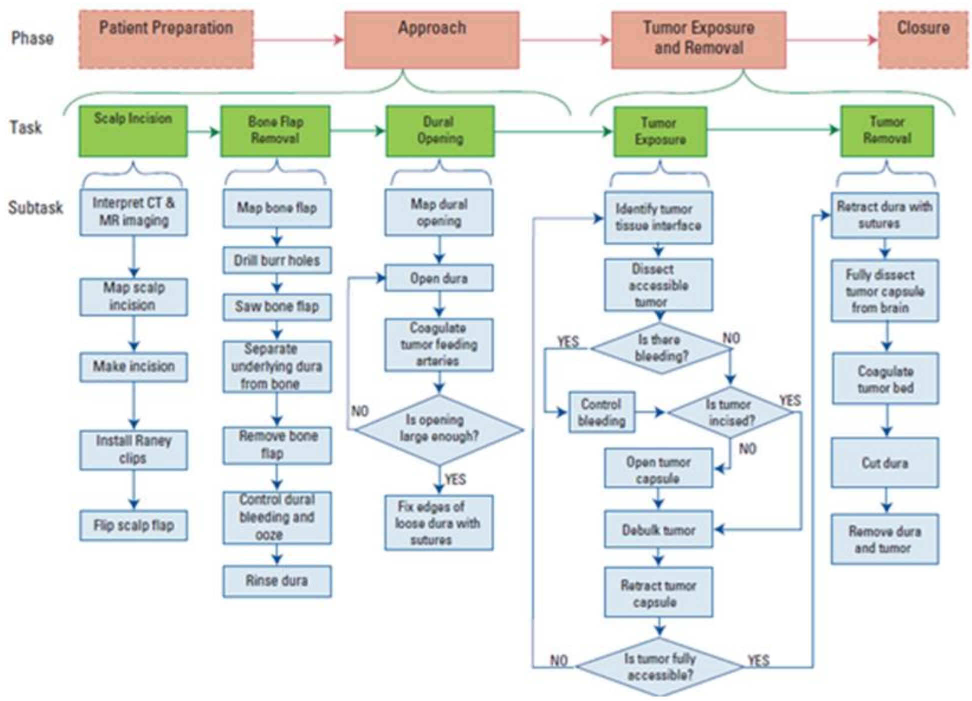
5.6.1 Cognitive Task Analysis

Both of these methods used in healthcare have had decades of successful implementation (i.e. Sarker, Chang, Albrani, & Vincent, 2008; Phipps, Meakin, Beatty, Nsoedo, & Parker, 2008; Demirel et al, 2016; Nagy, 1999). However, they have mainly provided descriptive information and cannot cater for cognitive components of the task under analysis. Mapping out HTA results alongside the cognitions of participants can provide additional depth and specific representative insight of each user's cognitions for every physical part of the system. To explore the knowledge requirements and *how* strategies are used to perform tasks, a method called Cognitive Task Analysis (CTA; Chipman, Schraagen, & Shalin, 2000) can be additionally implemented. This may be used as a tool to extract the cognitions of individuals at any skill level, to provide insight into their cognitive abilities during interaction within their system (see Figure 9 for an example of extracted decision-making). It can incorporate interview and observational strategies to produce a description of the knowledge that experts use to perform complex tasks. Essentially, complex tasks have been defined as the combination of both automated (strategic, unconscious, and procedural) and controlled (conscious, and conceptual) knowledge. The assumptions of CTA are based around the individual's cognitive abilities to create a series of simple yet intricate perceptual and cognitive processes for all performable tasks (Chapman & Sonnenberg, 2003). CTA can allow for the exploration and elicitation of participants' tacit and implicit knowledge as they provide meta-cognitive and reflective internal cognitive processes.

The individual's knowledge base, mental models, thought processes, and goal structures are captured by various interview and observational strategies (Chapman & Sonnenberg, 2003). For example, as one of the most frequently applied and prevailing knowledge elicitation tools observation can allow unobtrusive insight of CAP's whilst performing a task (Clarke, Feldon, van Merrienboer, Yates, & Early, 2008). Additionally, the expert can be probed into parts of the task through a knowledge audit and simulation interview to provide specific examples or information. Example knowledge categories elicited from an expert can surround perceptual skills, situation awareness, critical cues, recognizing anomalies, prediction, and improvisation. Once cognitive processes during the task have been elicited, a Cognitive Demands Table (Stanton et al, 2013) would be created for each task in the system/domain. Overall, the 5 common steps in most CTA methods are performed chronologically whereby the researcher would collect primary knowledge (with aid from HTA), identify knowledge representation, and apply focused knowledge elicitation methods (Chipman, Schraagen, & Shalin, 2000; Clark, Feldon, van Merrienboer, Yates, & Early, 2008; Coffey & Hoffman, 2003). Lastly, they would analyse and verify data acquired, and format results for the intended application.

Figure 9

A basic deconstructing the resection of a convexity meningioma. Note the inclusion of decision making for parts of the overall process (From Choudhury, Gelinas-Phaneuf, Delorme, & Maestro, 2013).



A specific variant of CTA named Applied Cognitive Task Analysis (ACTA; Militello & Hutton, 1998) included tools to produce information that can be more domain specific and cognitively relevant in nature. ACTA has been designed to further extract the critical cognitive strategies and demands associated with a task. ACTA has also been designed to analyse the cognitive strategies and demands associated with a task. Like a HTA diagram, a Task Diagram Interview (TDI) can be created that provides an in-depth overview of the task under analysis. However, a TDI emphasises the task elements that may be cognitively challenging.

5.7 Cognitive Task Analysis in Surgery

Such CTA variants have been used in highly technical fields including healthcare. Retention of technical skills in surgical procedures can be enhanced if CTA is applied on the procedure in question (Campbell et al, 2011; Luker, Sullivan, Peyre et al, 2008; Smink et al, 2012; Sullivan et al, 2007; Yates, Sullivan, & Clark, 2012). This is due to the exact identification of CAP’s cognitive processes (Schaafstal, Schraagen, & van Belro, 2000; Velmahos et al, 2004). Used with less experienced trainee surgeons, CTA can also reveal participants’ current understanding of the key procedural steps and surgical decisions (Flin, Youngson, & Yule, 2015), and thus identify any limitations in their understanding and knowledge.

5.8 General VS Surgery Specific NTS

From examining the current state of NTS research and training materials, one observation was made in this project that suggested that two types of NTS have been delineated with one being the prominent focus in training. To give labels, they may be called *general* or *surgery specific* NTS. General non-technical elements have been focused upon much more than surgery specific information. General NTS could be defined as cues and information which can be transferred to a wide variety of environments and are not unique to one type of surgery. Targeting the improvement of general cue identification has been the predominant perspective, upon reviewing current training content for the previously mentioned NTS courses/companies.

There has been benefit in that skills learned may be transferred to a variety of situations, however specific cues in each surgical procedure have been omitted and possibly unappreciated. Surgery specific cues relate to these very bottom levels of perceptual process within a Cognitive Task Analysis and provide information into the context and events within only one type of surgery. It may be argued that to help trainees become proficient in their specific surgical domain, unique cues and contextual information should also be included within training materials to improve the trainees’ abilities to identify surgery specific occurrences towards safety. This would disable the ability to distribute such materials to a wider audience but increase the skills of those who perform the specific procedure(s). By striking a balance of general and specific material a professional would have more flexibility to shift assessment from general events such as teamwork and communication protocol and technique, to specific phenomena such as cue identification within a technical task on a patient. Ideally, compiling a catalogue of surgical procedures overtime would create footage rich with surgery specific and general cues. These can be utilised to aid in both technical understanding and more in-depth feedback for an array of procedures and difficulties. Identification of general event should be utilised however there may be an optimal balance of information types.

Chapter 6 – Exploration of NTS in Surgical Trainees

6.1 Aims

Therefore, the aim of chapter 6 was to investigate if a homogenous cohort has a similar level of NTS, when grouped by parallel technical proficiency. This will allow understanding of where discrepancy may lie within a surgical training course. Responses from surgical trainees into their decision-making strategies in 3 written scenarios would demonstration differences in NTS can exist within and/or between year groups. The benefit of understanding these differences are key when comparing the data collected from the experts who train them, to find what information is required to help homogenise NTS in students.

Based on previous literature describing different decision-making strategies, it was predicted that 1st year individuals would more likely select rule-based answers as they would still be developing their perceptual and comprehensive abilities which limit dynamic and novel strategy. However, they may assign the issue to the supervisor either due to feeling lack of confidence in their ability (as suggested in Pulijala et al, 2018), or being overcautious to reduce the risk of error despite this delegation of duty. Conversely, 3rd year individuals were predicted to primarily select the analytical answer, which would indicate a more developed strategy to solve the defined issues. They should better understand rules and dynamic situations, therefore would be more likely to form a solution based on their refined understanding of the cues occurring in the situation

6.2 Participants

There were 28 students in each year for 3 years of their course. Participants were divided into 3 course years as this represented the technical skills proficiency of each year. Progression to the next stage at their respective institution allowed the assumption that an increase in technical competency had occurred, when compared to a previous year. Therefore, it was inferred that non-technical skills increased as the participants' course year increased.

6.3 Method

Three questions were included in a questionnaire given to the 85 postgraduate OMFS trainee participants within the 4 institutions visited for data collection in Pulijala, Ma, Pears, Peebles, and Ayoub (2018). These questions were asked to gauge the variation in NTS within individuals who have similar measures in technical skill proficiency. These questions provided an approximate measurement of their decision-making strategies and indicated if their answers were equivalent in the strategies used.

6.4 Materials

The three questions asked to the 85 participants were to find if there were differences in their non-technical judgements, whilst their technical skills were equivalent (See Appendix Figure 6). In question 1, the three answers each represented a certain decision-making technique. This meant that a rule-based, an analytical, and a recognition-primed option was provided (see section 3.4 for details of these constructs). Question 1 was: '*During a Le Fort I operation something unexpected happens. There is no procedure available. You have several seconds to make a decision to attempt to prevent a bad outcome, what is the best thing to do?'* Option 1 had an analytic underpinning, whereby all options should be considered before choice selection. Option 2 had a recognition-primed inference whereby the user immediately manages the issue in a more automated fashion, based on experience. Although an optimal choice for highly skilled individuals, there is more risk for error when less experienced individuals utilise this technique with different reasoning. Option 3 proposed the trainee to pass on the decision-making task to a superior to solve by use of their higher-level problem-solving ability and absent of information to apply a rule to the situation. The deciding cues in the question was '*you have several seconds to make a decision'*, as this inferred enough time for a more deliberate and thought-out approach. Also, a negative outcome would occur if a decision was not made.

In question 2, the same types of decision-making techniques were embedded in the answers however a context of communication and teamwork was provided. If decision-making theory and non-technical skills were varied within each group, this would provide an initial indicator to suggest possible inconsistency in participants' cognitive strategies. Question 3 was open-ended and asked, '*How do your senses help you when opening the Smith's Spreaders to downfracture the Maxilla?'* This question investigated each participant's perceptual experiences of performing this difficult task within a Le Fort 1 surgery. However, similar to the other 2 questions presented, the primary cognition-based aim was to measure the quantity of participants who can reflect, convert, and explicitly share their previous non-technical practices and experiences. Based on the prior notion that non-technical skills will increase with experience, it was predicted more of the 3rd year participant groups should be able to answer this question, and also with more details to *how* their senses help.

6.5 Results

For the questions presented to participants, descriptive and inferential statistics were performed using SPSS version 26 (IBM.com, 2019). A Kruskal-Wallis rank sum test was conducted to assess if there were significant differences in answers to the decision-making questions between the 3 different year groups. The Kruskal-Wallis test was used as it is a non-parametric alternative to the one-way ANOVA that does not share the ANOVA's distributional assumptions (Conover & Iman, 1981) and used for one ordinal independent variable with three levels and one ordinal dependant variable such as in the data.

Descriptive Statistics

The descriptive results for question 1 suggested a variation in choice selection both within each year and between the 3 years (see Table 1). There was a predominate split in choice selection for second and third-year participants between the rule-based and analytic options, and less choice to handover to supervisor. The third-year participants were less likely to ask for help from their supervisor/trainer (only 14%) than the other groups which partially concurred with the prediction that they would more likely form a solution based rather than handover. The descriptive results for question 2 showed 66% of all participants selected the option to '*interrupt the surgeon'* with 34% selecting the option to '*ask a team member'* if the information may be important. Although these 2 options were both plausible, the perspectives differed in that most considered more immediate and impactful behaviour was the better option towards patient safety (see Figure 10). The remainder selected the aid of their team to improve their understanding of the information to come to a better decision.

Table 1

Frequency Table showing each year groups question selections- there was a distribution both between and within the groups on their choices.

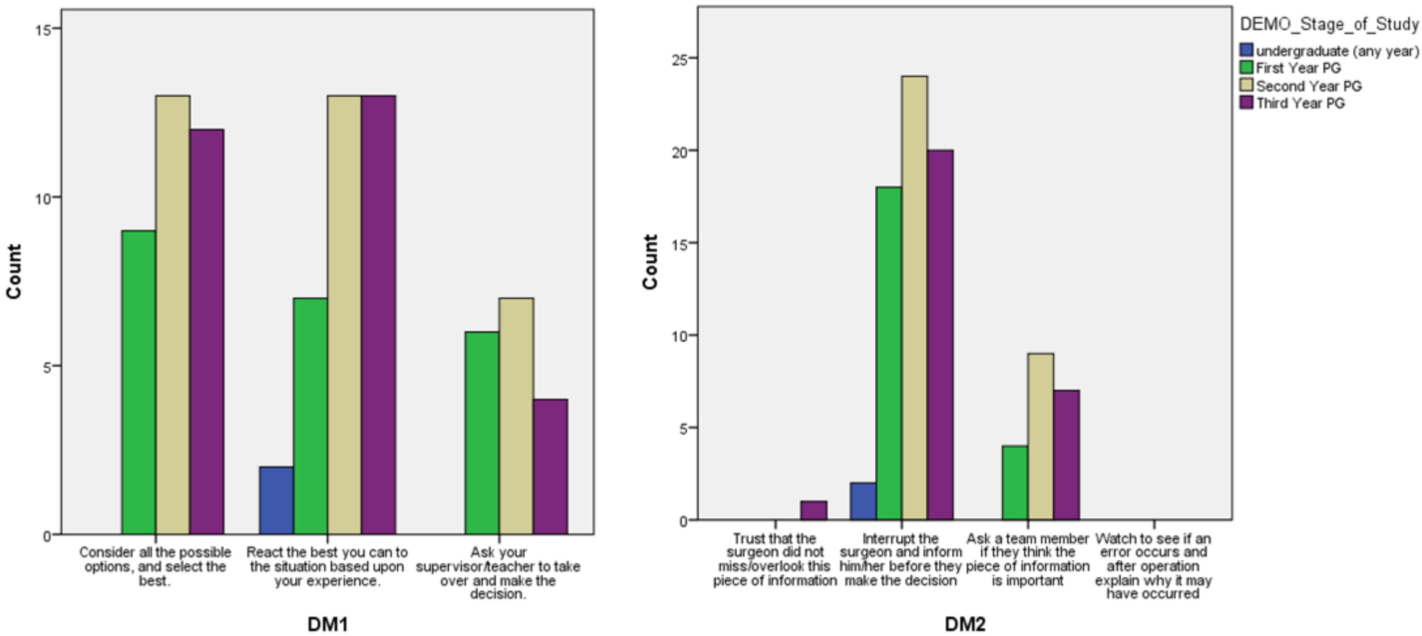
Questions

			Stage of Study			
			First Year	Second Year	Third Year	Total
DM1	Analytical Response	Count	9	13	12	34
		% within question 1	26.5%	38.2%	35.3%	100%
		% within Year of Study	37.5%	39.4%	41.4%	39.5%
		% of Total	10.5%	15.1%	14.0%	39.5%
	Recognition-Primed Response	Count	9	13	13	35
		% within question 1	25.7%	37.1%	37.1%	100%
		% within Stage of Study	37.5%	39.4%	44.8%	40.7%
		% of Total	10.5%	15.1%	15.1%	40.7%
	Rule-based handover of decision.	Count	6	7	4	17
		% within question 1	35.3%	41.2%	23.5%	100.0%
		% within Stage of Study	25%	21.2%	13.8%	19.8%
		% of Total	7.0%	8.1%	4.7%	19.8%
Total		Count	24	33	29	86
		% within question 1	27.9%	38.4%	33.7%	100%
		% within Stage of Study	100%	100%	100%	100%
		% of Total	27.9%	38.4%	33.7%	100%

Figure 10

Decision-making answers for 85 postgraduate surgeons. Different types of decision-making strategies were used within each year, and across all 3 years. This initially demonstrated how varied cognitive skills can be when technical skills were comparable

.



For question 3- '*How do your senses help you when opening the Smith's Spreaders to downfracture the Maxilla?* -five participants (3 in year 1, 2 in year 2) stated they could not answer as they have not performed the operation therefore 80 participants were analysed further. Thirty-two responses 8 answering with 'I don't know', 24 did not write any response. This meant these participants were not able to explain how they interpret their perceptual skills to help them perform the technical task of a downfracture, being separating two facial bones. In each year group there were 3-5 participants who had minimal descriptions (i.e. - '*feel the separation'*; 3rd year participant). More detailed description that had attempts of explanation with use of similes or other comparative language occurred 18 times for 3rd year participants, 15 times for 2nd year participants, and 7 times for first year participants. For example, '*I feel vibration along hard palate surface- similar to grating. An eggshell- like cracking effect. Hematoma submucosally and displaced maxilla' (1st year participant).*

Inferential Statistics

The results of the Kruskal-Wallis test were not significant based on an alpha value of 0.05. For question 1, $\chi2(2) = 0.46$, $p = .795$, indicating that the mean rank was similar for each level of Stage of study. For question 2, the results of the Kruskal-Wallis test were not significant based on an alpha value of 0.05, $\chi2(2) = 0.81$, $p = .667$, indicating that the mean rank of question 2 was similar for each Stage of study. No post-hoc pairwise comparison was required as these results indicate that there was not a year group that used different decision-making strategies more frequently than another year group.

6.6 Discussion

This abridged exploration of a surgical trainee cohort revealed that technical skills competence, as determined by successful progression of course, does not infer deviation of decision-making strategy from the previous year group (s). The results were mixed and opposes some previous literature showing differences in NTS usage for 3rd years compared to 1st years (e.g Purgh et al, 2014). Yet, Purge et al (2014) found no positive correlation between technical skills in intramuscular injection procedure, and non-technical skills, in the same set of participants. This difference in results is a key driver to increase effort into standardization of NTS training and assessment methods.

For question 2, the option to '*trust that the surgeon did not miss the information'* gained only 1 vote which suggested that either the hierarchical position of the surgeon was not a barrier for all other participants'- in that the surgeon's higher skill level was not to be overestimated. Instead, participants assumed it was possible to have missed information. This would concur with only a handful of research articles (Tsuburaya et al, 2016) that would interject the person regardless of status or experience. Although this action may change in a real-life OR, the results suggested that if there was a valid belief the individual can help improve the situation then they would interrupt a more experienced person. In question 3, the number of individuals who were able to reflect and communicate their experiences increased with each year. One explanation could surround the increased experience of third years with these perceptual events. Their cognitive abilities may have developed to a stage which allows awareness and feedback of their own actions in surgery, to reflect and improve upon their role within these events. This notion of self-reflection (e.g., Peshkepija et al, 2017) or introspection (see Ward, Wilson, Suss, Woody, & Hoffman, 2020) would need expansion to provide more reliable and replicable indicators, however not enough data was obtained (however it is explored later). Although a descriptive overview, these results highlighted the ranging abilities of each participant to explain how their NTS help them perform. Some participants had enhanced awareness of their tactile senses and cognitive events and were able to unpack and communicate these in some form- others were not able to do this. Additionally, for questions 1 and 2 the results generally suggested that participants used differing decision-making techniques when provided with the same information. There was limited opportunity to extrapolate this to more complex scenarios, however the underpinning phenomenon was that variation in non-technical skills to aid in behaviour was captured in this set of surgical trainees. This was not depending on the technical skills possessed as defined by their course year.

Chapter 7 - Utilizing Human Factors Methodology to Explore Cognition in Surgeons

7.1 Rationale and Aims

Subsequently, the aim of chapter 7 was to investigate and explore how tacit expert knowledge could be extracted and converted into training material for novices to aid in a standardised approach to training. The processes to complete this aim should have simple replicable stages for other researchers to follow as a framework thereby promoting extraction of expert knowledge into novel training packages. Within this chapter, the standpoint was not taken that technical and non-technical skills positively correlate as identified in chapter 6. The reviewed research suggested the strength of this correlation can vary for each individual and reliability of both skills being possesses simultaneously is inconsistent. Surgical trainees can possess stronger correlation of non-technical skills with technical proficiency while others have weaker correlation, even this could fluctuate over time. In general, it was posited that trainee participants' NTS would not be significantly different within the same course year, and that between course years there will not be significantly increased abilities of the more technically experienced group compared to their previous year(s). However, as literature suggested a significant difference, there was an expectation to observe different cognitive mechanisms used between CAPs and trainees participating in the experiment in Chapter 6. From this, the exploration of both general non-technical versus surgery specific material could guide the optimal structure for NTS training.

7.2 Method

7.2.1 Participants

Cognitively Advanced Practitioners were included in this project to help in the knowledge elicitation process. Three CAPs for interviewing were necessary due to each of their limitation in explanations from automaticity from practice. By including two or more CAPs, the data gathered from each person can be combined. An analogy of this may be if a driver in a car cannot view certain parts of the road behind them (etymology of *blind spots*), then the inclusion of one or more passengers can form a more inclusive perspective to minimise the issues caused from blind spots.Three participating CAPs were recruited in India from Goa Dental College-Goa, Nair Hospital Dental College-Mumbai, SDM college of Dental Sciences-Dharwad, and Manipal College of Dental Sciences-Mangalore. This was performed parallel with the data collection in Pulijala, Ma, Pears, Peebles, and Ayoub (2018). There were 5 initial participants in total however 2 were not considered Cognitively Advanced Practitioners after reviewing their interviews as they did not meet the criteria in section 5.3 in addition to repeating a portion of information covered by the 3 main participants. The participants were primarily selected due to the increased availability of consultants and Head of Departments when compared to the U.K.

Participant 1 was the Head of Department at one of the institutions. She has had a total of 20-22 years of surgical experience with more than half of those years teaching others. She first published research in 1987 and has published over 40 articles. She regularly trained students in surgery. Participant 2 was an undergraduate surgical student in 1986 and postgraduate at the same institution. Staying in the same location up to the point of interview, participant 2 regularly taught 2 to 3 trainees per operation and has been specifically teaching Le Fort 1 Osteotomies for 15 to 18 years. His overall surgical experience after undergraduate training was 27 years up to 2017. He has published 47 research items in OMFS surgery. Participant 2 had never been explicitly taught non-technical skills with the definitions used in this project- he stated they were gained as he '*just had experiences on my own'* over his career (L20). Participant 3 was Professor and Head of Department for Oral and Maxillofacial Surgery. The researcher identified participant 3 to most likely have cognitive abilities that would be categorised at a master level, within the Apprenticeship Model (Dreyfus & Dreyfus, 1996). Participant 3 stated that he helped to build a small part of the college in 1980 when it opened and has worked there since. He has been a staff member practicing OMFS surgery for 30+ years and at the date of interview still teaches in the OT 2-3 times per week. Participant 3 had 28 research publications. The intensity and frequency of surgeries for this long time period, along with the frequency of training novices also made participant 3 the comparable candidate most likely able to make explicit and verbalize his cognitions in a form that trainees can understand.

7.3 Design

One-hour semi-structured cognitive interviews were given to participants separately. These interviews gained richly detailed insight of first-hand experience. The non-probabilistic method of purposive sampling was used to specifically seek participants that fell within the category of *expert* driven by the Apprenticeship Model (Dreyfus & Dreyfus, 1996) and Human Factors perspective. This meant that each criterion of career analysis, sociometric analysis, and performance analysis criteria (Crispen & Hoffman, 2016) was initially used to sort and filter the staff in each institution until more details could be gained from the interviews themselves. Then initial probing questions were able to more accurately determine if a staff member had significantly different cognitive capabilities when compared to their peers. The design had flexibility in that the interview questions were created in a structured format so that depending upon the participant's answer, pre-planned routes could be taken which probe deeper into the direction taken. This structure enabled the narrowing of the tasks in question until underpinning details of the participant's cognition were then explored with novel and more specific probing. This design accommodated for when participants' responses affected how and which questions were asked next, as they were predefined.

7.4 Materials

7.4.1 Task Analysis

An initial Task Analysis for a complex Le Fort 1 Osteotomy was created before interview, for the researcher to better understand the topic. This was chosen as it was following from the awareness provided by Coffey-Zern, Calvi, Vorrasi, and Meara (2015) (see section 4.03) of safety-orientated training shortage in OMFS. Cognitive exploration and Task Analysis in this department have been less common than other surgical areas yet are necessary to improve training. As the consultant surgeon in OMFS surgery (co-supervisor) in the UK had limited availability, much of the TA was created by the merger of other media sources to the researcher's best understanding. This saved time when interacting with the surgeon to be able to clarify the steps. The goal was not to produce a TA which could be used as educational material for students which details each step, but to capture the depth and composition of each task to be able to later add the cognitive events associated with each task. Materials used to create the TA were popular handbooks (Ayoub, Khambay, & Benington, 2013; Haggerty & Laughin, 2015), Online video (Jacquemart, 2014), and a leading mobile application (aofoundation.org, 2019).

7.4.2 Cognitive Task Analysis

After performing a TA, a researcher can cognitively interview a CAP to allow the exploration and elicitation of their tacit and implicit knowledge, as they provide meta-cognitive and reflective internal cognitive processes. The researcher identified Cognitive Task Analysis (CTA; Chipman, Schraagen, & Shalin, 2000) as the most appropriate knowledge elicitation technique as it can expose cognitions in experts and trainee's (Flin, Youngson, & Yule, 2015). This is the content in the semi structured interviews that took place. Secondly, each participant's cognitions can be mapped out to provide a representative, in-depth insight of their technical and NTS. Lastly, this method is streamlined and not resource intensive compare to other techniques that have developed in other domains (see Militello & Hutton, 2000). As a surgeon practices many procedures over long time periods, they rely on increased automatic decision–making models and use of compressed information. This method within the repertoire of Human Factors to decompress such knowledge was chosen as it has been successful in extracting the reason behind underpinning behaviour, in both trainees and senior surgeons (Jaffe et al, 2017).

A predefined list of human factors-based questions was created for the CAPs from the Task Analysis performed before data collection in India. These aided in the decompression of the CAP's expertise as they were deeper probing questions that adapted to each answer, the resultant output becomes the final Cognitive Task Analysis. A check for leading, ambiguous, or assumptive questions occurred to ensure no inadvertent biases in responses. Orienting around the *how* and *why,* these questions were designed to stimulate further reflection in the CAPs cognitive processes. Some examples were- 'Why *would you instruct to tap the mallet twice, but tap it only once on the bone yourself?', 'How do you accommodate for the prevention of undesired events when teaching?'* (See Appendix Figure 7 for a list of predefined questions used).

Typical terminology used in the academic literature was changed after a trial interview (not used in this thesis) as the terms significantly limited the depth the researcher was able to probe. The main terms were '*problem*', '*error*', and *'issue'*- when used within questions there was circumvention on specific events. New questions were designed and typically the term 'undesired event(s)' was used which was much more receptive in subsequent interviews. All words were revisited and adjusted to also infer an undesired or unpredicted change rather than being attributable phrases. However, the effects of the semantic changes were considered in the analysis.

7.5 Procedure

Probing questions were designed around certain parts of the surgery which were more technically demanding, or if there was suggestion of high cognitive workload (see Appendix Figure 7). A general understanding of the tools, anatomy, and procedural steps was also learnt for the researcher to be able to understand the interview context better. The starting point for the TA began with an observation of a Le Fort 1 Osteotomy in surgery with notetaking. This provided an opportunity to observe the theatre and the way in which the individuals differed in cognitive and technical skills. The benefit of this dynamic interaction was that there were verbal instructions, questions, and answers, representing transfer of knowledge/cognitions/ perspectives. Written notes were taken to retain the order of the tasks and provided the ability to transcribe and add notes, to further the understanding of the surgeon and the task under investigation. Questions were asked with underpinning relevance to psychological theory. These answers helped the researcher understand the perspective of the surgical team.

Then, funding for 3 weeks was provided by the successful application to the Santander Mobility Fund Award (Santander.com, 2018) in support of the project. The four universities were contacted, and the heads of departments arranged to meet the research team, along with a tour of facilities. After finding out the staff members and their roles the researcher approached individuals who were most likely to be of a higher level in the Apprenticeship Model. Upon meeting each participant and being interviewed singly, the relevant information sheets and consent forms were presented (see Appendix Figure 10). Full disclosure of the interview was given, and participants had an opportunity to ask any questions and/or decline participation.

There has been caution with use of cognitive interviews to allow participants with time to process the questions asked (Stanton et al, 2013). It can be difficult to explain automated action and participants could react negatively to such probing techniques if pressure is applied for a response. Therefore, a slow pace was given to interviews with a friendly and open manner. Participants could ask to move on without question. One participant showed frustration in the inability to explain an answer, but later answered the question once they had thought about it more. Although not presented in this project the researcher asked some of the trainees present about their mentors to understand their level of agreement regarding of sections pertaining to the communication of the mentors with trainees. They concurred with the answers provided by the CAPs (i.e., '*Yes, he does ask us for our opinion').* The interviews were transcribed, and specific information was collated for the Task Analysis to produce a detailed description which showed the cognitive elements extracted from the participants for some of the technically challenging tasks in Le Fort 1 surgery. The other non-technical elements extracted were analysed separately and consequently presented.

7.6 Thematic Analysis Process

The data were analysed by thematic analysis (TA) as informed by Braun and Clarkes approach to data collection and analysis (Braun et al, 2015). With a research aim in mind and of approximate potential findings, this TA was firmly a deductive approach- in that existing theoretical concepts provided a foundation for how to 'see' and provide 'meaning' to the codes, then clustered into the final developed themes (Braun et al, 2015). For theme development the familiarization of examination codes, and combining or clustering, and subsequently collapsing codes can aid in forming meaningful patterns. To start this process, richer and more complex code were extracted and examined to understand what other codes were within them, these can then be formed into an initial rough list of themes (Charmaz, 2000).

However, more considerations are required to elicit the features of similarity and relationship between codes which may appear to have majority differences. There can be difficulty to perform this process until the absolute central organising concept has been identified (Braun, Clarke, & Terry, 2013). This has difficulty as insufficient effort to extract themes, addition of biases, or misinterpretations of data can obscure the absolute central organising concept being realised within a theme.

Once identified however, new code can be compared to assert whether it fits well within the theme or not. This process will change many times during the analysis process as the researcher's understanding of what they are forming from the code, also develops. This method is informally called *patterning* of themes. As the researcher builds upon their initial understanding of the raw text, they engage in selection of relevant segments of data, which allows them to find the implicit patterns occurring based on the research questions in mind. Ultimately, patterns should start to occur in the data, and was experienced when analysing data in this thesis.

A thematic map produced at this early stage will not be the optimal representation of the themes within the data. Instead, use of a rough visual map can help to apply productive and reflective data engagement techniques. Boundaries of themes can be better understood, and their potential relationships plotted on such thematic maps or tables (Braun, Clarke, & Terry, 2013) help to further understand node relationships. Once initial patterning of data segments and relational themes (with reference to each central organising concepts), have been better solidified the next step is defining and naming of themes.

To achieve this, summation of themes (being to be analytical in the data structures and codes) should be switched to interpretation of meanings. In this thesis, this aided in writing and developing the narrative that is encased in each of the themes formed. A benefit of approach is that when difficulty occurs to create a clear and cohesive narrative, this likely indicates the data is still to be developed. Data that skews interpretation away from the central organising concept can be better identified when preventing the researcher to write the narrative of each theme.

This was the general method used however this is not strictly required from TA. A strength of TA in the application towards insightful findings is also a weakness in its methodology. This property is that there is not particular set method to follow. Indeed, there are recommendations for the protocol (such as Braun & Clarke, 2006, p96) but its openness to adaption allows any researcher to use their own means to identify patterns and form conclusions. For example, TA was used with such controls as recording reflective notes and collaborative analysis, to explore implications of COVI-19 on surgeons (Whelehan, Connelly, & Ridway, 2020). Like many other methodologies pattern identification is a key aspect in their goal. However, the choice to use TA lies in its dependency on the research- to be successful in producing in-depth analysis by any interpretations required by the data within a project. The malleable platform depends upon production of a defensible analysis due to effortful researcher engagement, moreover methodologies that hold more constraints that may constrict justification of a process- but only to stay within its boundaries.

7.7 Results

7.7.1 Task Analysis

The TA consisted of 28 main steps, which broke down into more than 200 basal tasks/actions. The information categories were structured by: Main Task, Sub-Task, Basal Actions, Task Description, Goal, Time Taken, Equipment Used, Task Difficulty, Potential Errors, Meaning behind Action, and Type of Process. There was a substantial amount of information to form further cognitive interview questions regarding surgical tasks and the dynamic events in the OR. A full version of the TA can be viewed upon request.

7.7.2 CAP's Cognitive Interview Findings

With this basal understanding of how non-technical skills were distributed, the CAPs were then interviewed which would allow insight into how they utilise their skills to increase patient safety. See Appendix Figures 8A-

8C for interview transcripts and analysis segments. Interviewing the CAPs also allowed exploration into how their advanced NTS levels have been used to adapt when teaching in the OR. It uncovered the cognitive adaptations they implement that are required to maintain safety against error and demonstrated how important real-life guidance by a CAP can be. These expert-trainee differences presented in this chapter may not have been new within Human Factors research, but it was the assessment and transformation of the phenomena uncovered within the specific context of surgical training and VR that possessed novel value. The results and brief discussion presented in this chapter determined which advanced NTS elements could be replicated outside of the OR to address problematic events, with use of customized teaching techniques within a virtual reality platform.

The NTS related phenomenon were presented in order of the extent to which they were embedded within their interviews. This meant that the more explicitly mentioned and explained NTS related processes were presented first. Then, more in-depth phenomena were presented that may not have been explicitly stated but were extracted with reference to appropriate cognitive or surgical theory, and human factors techniques. The NTS were then split into 2 overarching categories- those that were only applicable to the surgery type discussed (Le Fort 1; called 'surgery specific skills'), and those that were able to be transferred to other surgeries and situations (called 'General NTS' for this project).

7.7.21 Purposeful augmentation of NTS as a countermeasure when teaching

(Key NTS Verbalized**:** Shared Situational Awareness; Shared Mental Model; Projection/Prediction; Communication; Teamwork)

Each CAP identified and described the implementation of their non-technical skills in an enhanced manner under teaching circumstances, as an additional safety factor to maintain patient-safety when instructing surgical trainees. When in training, the CAPs compensated for the limitations in their students' NTS by temporarily augmenting their own. However, they suggested more cognitive effort is then expended and their abilities can sometimes be compromised when such added requirements are placed upon them. To illustrate this technique and specific components captured, the comparison of participants' actions when in a non-teaching situation was firstly presented:

In a non-teaching environment, when describing their team's communicative and predictive abilities CAP 2 stated that if each team member has a sufficient level of predictive powers, he can save his cognitive expenditure as these tasks can be distributed between each individual. He stated*,*

'because we are doing a large volume of cases…. they already have an idea about what I ask for next, so they come prepared. So not much of a talk really happens…They know what I will ask for and they get it immediately' (CAP1: L306).

Participants 1 and 3 also described how communication and decision-making can be simpler and optimised when working with similarly skilled colleagues. They stated,

'*They know what I'm doing, they know what experience I have, and they know what, erm, what fears I have'* (CAP1: L221). '*it gives a lot of confidence when you are doing surgery. You are going with the team. And you are all clear of your concepts and you are taking everybody with you'* (CAP3: L233).

These views indicated that if participants work with experienced people, they communicated minimally as they can infer each other's judgments and they understand the premise of each other's thoughts. This can be described as the gradual conversion of a team's conceptual and knowledge related information to form shared mental models (*see Discussion*), therefore synchronised to what is important (Van den Bossche, Gijselaers, Segers, Woltjer, & Kirschner, 2011). Comparably when participants teach trainee surgeons, adaptations of their NTS were used to account for this change in team abilities. Participant 1 stated that she frequently reminded students of certain high-risk cues as they are doing a task (CAP1: l143). Being asked how she works with trainees, CAP 1 stated,

'*I would have to totally take charge and totally be highly alert to how they are behaving on table and how they are handling instruments, because it gives a little bit of tension when you don't have other experienced people around you*' (CAP1: l231).

This effortful behaviour of additional vigilance is a proactive action deployed to counteract the effect from the loss of colleagues, who both had increased levels of NTS and technical competence. Similarly, proactive augmentation of NTS for CAP2 was identified when he explained that he knew that trainees will not have the predictive abilities to avoid an undesired event, or to assess solutions and make effective decisions if such events occur. When CAP 2 recalled training individuals (see Appendix Figure 8A) he described them to be within the first 2 levels of situational awareness, in that they are still comprehending the core elements and interactions in the system. CAP stated he would have to '*think ahead'* for:

'...the assisting surgeon, and the nurse- the OT nurse' (CAP2: l325). '*if they have had less experience you have to keep speaking, and for these two you think ahead for' (CAP2: l330).*

Furthermore, his socially orientated NTS workload would be typically much higher than usual, as he described that:

'not much talk happens. If I had to do a same thing in a theatre apart from this hospital, it would involve a lot of talk'... 'Because they don't have an idea of what I'm asking and what's next so I will tell them' (CAP2: l310, l322). In alignment with this, participant 3 stated that '*More effort sure, more effort'* (CAP3: l245), is used.

CAP 2 developed this point to state how he would go into the theatre before starting the operations and evaluated the environment. This action suggested that CAP2 attempted to build up his mental model of the new operating theatre, the layout, and the available tools to improve his decision-making abilities during surgery. During surgery he detailed how he can increase his projection to help ease an undesired future situation. E.g. he stated,

'I need to see that the equipment is kept ready, then in the OR I imagine a little further and then I tell them 'ok keep a roller gauze pack, and get the Surgicel available to me'' (CAP2: l316).

These proactive behaviours demonstrated the participants' understandings of the limitations of NTS within trainees, and how they depend on and adjust their NTS to compensate. Moreover, this would not be done if the participants were with individuals who were able to distribute situational awareness on a higher, projective level. The consequence of this can be varied, however for CAP 1 she described how trainees can lose concentration on a task, which could further impact their understanding of the situation in a live point-of-time, therefore adding more reliance on the CAP to counteract this behaviour. She stated, *'So, there might be a little bit of shouting and screaming at them if I find that they are not onto it, so I have to erm, be watchful not only of what I am doing, but how they are assisting me* (CAP1: l122)'.

However, although each of the three CAPs interviewed were explicitly aware of how they adapt their NTS when teaching, another phenomenon was uncovered which was not expected in the project- which suggested they attempted to use technical adaptations in parallel. Instead of augmenting their NTS as a primary measure to maintain patient safety when training, they described how they had adjusted several technical tasks to bypass the majority of common issues trainees may produce. The results created a reduction in cognitive load (*see section 7.31*) and reduction of error in theatre, and were called '*Compensatory Behaviours and Instructional Adaptation'*. These were classed as 'Surgery Specific' non-technical skills.

7.7.22 Compensatory Behaviours and Instructional Adaptation

(Key NTS Extracted: Projection/Prediction; Leadership; Decision-Making; Communication; Situational Awareness; Mental Models)

Due to experiencing and managing common errors first-hand over several combined decades of full-time teaching, each CAP had formed new ways to better adapt tasks to achieve the outcome required. The researcher considered this to be a mechanism for the participants to reduce their effortful augmentation of NTS described in the previous section. For example, participant 1 stated she changes her technical behaviour to counter overenthusiasm,

'*they are very eager to pull on the retractor. So, to be on the careful side, I generally keep my incision a little short of the molar…because I know that at times students are too enthusiastic* (l135)'.

P3 also adapted a difficult technical task by providing the trainees with a less efficiently performing tool (a saw to cut bone). The effects of this were that the trainees are forced to slow down their action, as the saw cuts slower than the typical saw used. Additionally, if a trainee starts to perform the task incorrectly, the participant has more time to identify, prevent, and explain the issue to a trainee, as the saw would not damage tissue at a fast rate. Likewise, although CAP2 performs a specific action in one attempt ('*in one tap of the Mallet'*), he had predicted the actions of trainees and alters instruction to rectify their predicted error:

'*I see the trainees when they put one single tap, they tend to use a lot of force. You know, so it goes a little uncontrolled. So, when I tell them to do two taps they tend to use the force in a more cautious manner. So, that probably is the reason why I tell them to do two taps…'*

This suggested that the trainees' cognitive abilities may not be at a stage that permits synergy with their technical abilities. I.e. although their 'hands-on' skills may allow them to complete the task, their approach may not have the ideal characteristics under certain circumstances. There may be possibility that non-technical training could help individuals be more aware of their thoughts and actions to reduce the occurrences of being 'eager', 'too-enthusiastic', and 'uncontrolled' as described. This could be fulfilled by contrasting and presenting the mechanisms used by the CAPs in the same situation for trainees to understand this typical limitation. There were several examples of this phenomenon, and the cognitive interview technique was crucial in probing into the reasons behind the CAP's cognitions and behaviour. This provided a small insight into how complex surgical training in the OR can be, and how fluctuations in NTS can alter the dynamics and experiences for each individual present.

7.7.23 Decision-Making Towards Time Management – *Buying Time to Think*

This theme was much more embedded within the transcripts- CAP2 stated that when an undesired event occurred, if bleeding could not be sourced from a visible location then the only option is to find the source of the bleeding by applying pressure (typical technique). He described his goal to find the source of bleeding. For example, he stated, '*I tell the anaesthetist there is a bleed and erm, they say 'hold on' and they try to get the blood pressure down* (CAP2: l220)'.

Furthermore, he stated, '*there are multiple areas it can bleed from, so I would rather advise them to put a pack and hold on for 5 minutes… to think it over a little bit* (CAP2; l242) *If you're not sure what is happening, then you get surgicel* (a gel to make blood coagulate) *…wait for ten minutes* (CAP2: l256)*'.*

However, in describing each process his underpinning decision-making processes at a more basal level were extracted. This additional goal was embedded within his description of rectifying a bleeding related issue. This goal was to decide how to slow down the bleeding to generate more time to assess the new situation. Through slowing down the rate of bleeding, this allowed for a wider range of non-technical abilities to be implemented, and in a more controlled method. Therefore, a secondary aim of 'packing' blood was to have better control of time as this allowed the team to collect new information before forming a new recognition-primed plan. This was a prerequisite to then solve the main goal of locating and stopping the bleeding in a better manner than on the immediate recognition-primed response.

He described that a subsequent set of cognitive processes could be performed *only after* he has controlled the amount of time, which was fulfilled by various 'packing' methods.

For example, CAP2 stated *'…get it under control. And then start thinking about it. Whether my osteotomy cut was higher, 'did I do the disjunction correctly', when I was retracting the soft tissue* (l265).

In this next cognitive stage to find the source of an undesired event, CAP2 described how he recalls his actions to review the most likely cause, and most likely location of the issue. This could be classed as use of a higher-level 'Reconciliation Cycle' as defined by Cristancho et al (2016). Perhaps viewed as a situational awareness level 3.5, the CAP had transformed previous cues to provide different information to improve his mental model. He was able to reflect upon his input into the system (surgery) because of this and understand both the short- and long-term effects on the patient. This weighing up and flexibility in predictive power are the two components of the 'Reconciliation Cycle'. He stated, '*Yeah, I go back to go forward (*l273*)*' which supported the use of this strategy as it demonstrated a refinement of knowledge to create additional information not previously existing. Therefore, CAP2 demonstrated a high-level cognitive ability to control time and readjust current understanding by transforming previous information, thus re-establishing a more accurate projective (level 3) situational awareness after unpredicted change had occurred.

7.8 Discussion

Based on the complexities of the *surgery specific* results, there may be difficulty in accurately transforming the information extracted into training material. Several decision-points emerged that oriented around surgeons' abilities to make checks before performing a high-risk tasks, cues/sensory information associated with such a task, and decision of how best to communicate actions prior/during/after error occurrence within these tasks. The context and content of these cognitions were only applicable to Le Fort 1 and as such the skills of the researcher were not sufficient enough to accurately extract these cognitions from a healthcare perspective. Therefore, these were deemed beyond the skills and comprehension of the researcher and not attempted to be made into training material.

Moreover, as described in 6.72 '*Compensatory Behaviours and Instructional Adaptation*', the experts compensate their behaviours to support the trainees, and in doing so optimise their desired outcome. There may be severe difficulty re-enacting this in VR and although users can follow instructions, it would be difficult to provide an experience that replicates the OR because of these modified behaviours of the experts. They would be difficult to transfer through digital media. This may be the limit to which the depth of digital NTS scenarios can provide- in that decision-making on a micro scale can be shown and trained (i.e. when/where to stop cutting), but the transfer for such learned behaviours into real life may not match. As experienced tutors will modify practice in unique and dynamic styles for each trainee. This would result in trainees either unlearning previous information with great difficulty, or still being required to learn more skills regardless of VR usage.

However, there were cognitive mechanisms and behaviours which were demonstrated to be important for the CAPs to perform safely. The interview results allowed the researcher to capture the underpinning thought processes of each CAP, and with reference to cognitive theory it may be possible- to a degree yet unidentified in the research area- to explicate these processes in a simulated environment. To expand on this point, the results in the theme '*6.84 Decision-Making Towards Time Management – Buying Time to Think'*, can be matched to the 4 different types of decision making (recognition primed, rule-based, analytical, and creative; *see section 3.42*). The theme described the experts using recognition primed processes to establish bleeding, make rule-based decisions of how to slow down bleeding, then shift to analytical considerations as more time has been created to better assess the situation, with final decision of how to rectify the issue which could be creative where required. Upon reviewing the interview transcripts, there was previous theoretical support for each of the expert's explanation and were discussed later in this project. Superficially, this may appear to support these interview results to concur with other expert as whose NTS also have already been capture by research. However, understanding the surgical processes and capturing surgical cognition allowed a more realistic identification of what NTS can be transformed into digital content, how this may occur, and what impact/usefulness this would have on trainees.

Chapter 8 – Psychological and Educational Considerations for NTS Investigation

8.0 Aims of Creating the Non-Technical Skills Training Material

This chapter reviews and combines both the pragmatic elements of results found in chapter 6 that was postulated to be added into the experimental VR trainer, along with assessing the relevant theoretical and practical foundations of education and psychology that also need to be incorporated. The educational comparison of the VR trainer to current material was essential for future usage as no recognised formal theoretical framework on how to teach NTS has yet to be implemented (Bourrier, 2016). In surgical simulation research (Kleven et al, 2014; Mathur, 2015, see also Ahmed & Dan, 2015) there has not been an explicit framework or guiding principles for the application of Human Factors assessment, improvement, and re-evaluation of VR systems. However, the previous literature review has provided support for the necessity of this framework to lay detailed foundations and guidelines if the increases of other VR related cognitive studies occurs. The goal of this or a similar program, would be to enhance the perceptual and cue-recognition abilities of the user. Thus, facilitating higher levels of cognitive skillset primarily aided by expert feedback.

Like a teacher, an application should not provide direct instructions to a trainee but instead direct their attention to the area that contain features relevant to the trainee's goal (Sheehan & Higgs, 2013). A cognition-based trainer should have the properties to be able to direct the learner's attention on the salient task features and identify consequences of actions/risks for differing scenarios. Guidance should be deliberately provided but the structure and nature of the scaffolding process infers trainees should have effortful experiences with minimal amount of help when performing scenarios. If or when failure occurs, insightful information can then be added as clues and small amounts of new information can help scaffold the user's interpretations with deeper understanding than provided by simple instruction alone.

8.0.1 Further Theoretical Considerations

Considering the characteristics and properties of one-to-one training in the Operating Theatre (being the gold standard) typical coaching and training models were required to be incorporated into the application. This was necessary to attempt to produce a simulation which can be close to, or match the superior elements found in OR training.

By comparing traditional surgical coaching setup, adaptations were made to these formats that could also be as effective in a digital format. Surgical coaching can allow feedback from a surgical expert who can observe a trainee and objectively explain what can be improved (Greenberg & Klingensmith, 2015). This formative feedback can help a trainee reflect upon their cognitions along with their surgical actions and critical self-awareness can be encouraged. The resultant improvements should be a trainee with better understanding of their own cognitions (metacognition), and how to modify their cognitions to better match the expert's explanations (Bonrath, Dedy, Gordon, & Grantcharov, 2015). They also may improve their reflective abilities from their actions in future situations. From this, Collins, Brown, and Holum (1991) proposed the Cognitive Apprenticeship Model which posited that throughout the intraoperative process there were 6 methods that are desired for excellent clinical teaching. In order, these are: modelling, scaffolding, coaching, articulation, reflection, and exploration. Linearly completing these methods, the model dictated that a surgical expert should be able to explain tacit and implicit steps - describing them in a manner understandable to the receiver (i.e. the learner). This model has had success in demanding environments such as emergency medicine (Merritt et al, 2018) which suggested it is also possible in other healthcare situations. The resultant skills learned by the apprentice should have both specific procedural cognitive skills but also general understanding that can be transferred and generalized to other areas. This model is expended in the discussion.

8.1 Further Practical Considerations

8.1.1 User's Cognitive Load for Optimal Experience

The integration of these 6 stages of coaching was established as an objective when designing the training material. However, as participants will not have used the resulting program, three combined novelties performed in parallel may create difficulty when used in the short-term experimental setup of this project. Within a limited timeframe, users will have to actively engage to improve their practical NTS, as well as interacting with a new VR technology, and a new interface. Users may be overwhelmed due to such high rate of incoming information and adaption to the system, in that modelling, scaffolding, coaching, articulation, reflection, and exploration may not efficiently occur. To counter this, Cognitive Load Theory (Miller,1956; Sweller, 1988) was explored (see Figure 11 for a representation). It is one of the prominent theories of pedagogy and provided a theoretical framework of cognitive architecture. The main premise is that working memory and information processing is limited. With increasing consideration in medical simulation (Anderson et al, 2016), the theory suggests three sources of cognitive load in any learning situation. The fundamental load of the learning task (being surgical training), the extraneous load of the learning situation (being the VR system), and the cognitive load required for the learning process to happen. Learning and skills acquisition can be hindered if a user's total cognitive load is not enough for the demands of the learning system. The limits in working memory and processing of information are exceeded as the learning tasks overexpose an individual to extraneous cognitive load e.g., when the learner needs to integrate complex multiple sources of information (Kirschner, Sweller, Kirscher, & Zambrano, 2018)

For users to gain positive experience and increase likelihood of additional practice, users' cognitive load needed consideration in the learning environment. As this project used low-fidelity simulation (compared to real-life simulations), it could be hypothesised that VR can facilitate a better learning experience as it can better control, analyse, and present complex events when compared to real tasks (see Pulijala et al, 2018). Novel visual and auditory cues could also impose additional cognitive loads compared with other training modalities. Therefore, simplified VR training should be suited for novices who have more likelihood of encountering cognitive overload. Increasing amounts of information present should only increase with user's proficiency to process more information.

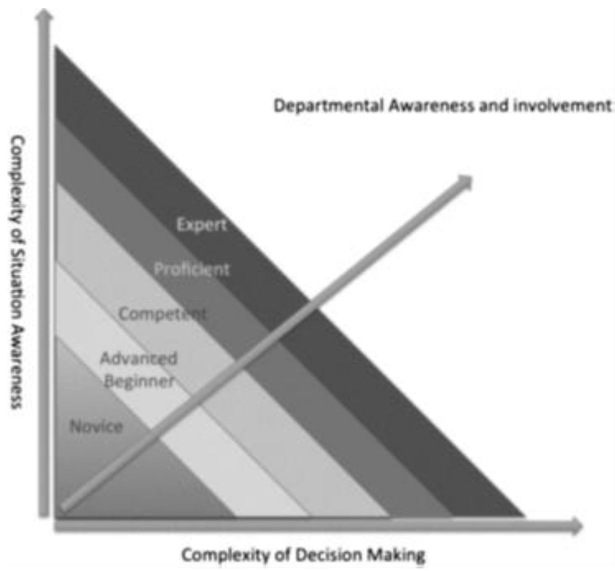


Figure 11: This figure shows the general direction and complexity of decision making with situational awareness. (From Lowe et al, 2016).

8.2 Consideration of Measures of Cognition-Based Assessment

There has been over 30 widely used human performance measures and assessments matrices (e.g. the SESAR project). For this project several methods were identified as relevant to capture the cognitive processes of participants and were reviewed before final selection for incorporation into the application.

To measure real time performance as an indicator of SA, the Situation Present Assessment Method (SPAM; Durso et al., 1998; Endsley, 1988) was formed and is an atypical situational awareness tool as it does not remove the user from the situation. The SPAM has demonstrated its ability to distinguish between expertise differences in chess (Cruchfiel, Moertl, Ohrt, 1995), automation failures in air traffic control simulations, and individual differences in air traffic control trainee potential (Gronlund, Dougherty, Durso, & Canning, 2001). The SPAM technique provides assessment of mental workload and situational awareness at the same time. The SPAM can be easy to administer and requires minimal training. It provides an objective measure of situational awareness. There is no need for task freezes which can distract and disrupt users therefore administering probes in real-time allows the technique to be applied 'in the field'. To measure workload, each SPAM query initializes with a warning signal, to which the participant must acknowledge. The SPAM would continue once the participant has acknowledged the signal. If the participant's momentary workload is high, the expectation is that more time will pass between first presentation of the warning symbol and the participants acknowledgement of it.

The rationale behind utilizing latency as a measure is that if participants have strong situational awareness skills, they will either seek necessary information to answer the question or will already have the knowledge in memory. Either way the latency should be minimal. Alternatively, if the situational awareness of the participant is low then there will be more inefficiency within the information gathering process. This is because they may need to find more information as they have not gathered the same amount as an individual with strong situational awareness skills. Also, their information gathering methodology may be more inefficient therefore delaying response time. The participant continues to engage in the task, and if they are not able to successfully attend to the request of the SPAM based upon their current involved working memory, the situation can be assessed to determine the possible factors disabling the user from attending to the SPAM. The dependent measures from the SPAM are the correctness of participant's response, and the latency to generate the response.

However, various preparatory activities are required, including the conduct of situational awareness requirements analysis,. Even if task freezes is not applied the method is intrusive to task performance. Attention may be redirected from the required situational awareness elements therefore affect the measures of the user, due to its own interactions. The SPAM was a possible measure to use however a requirement was that it did not distract the participant during a task. A similar measure that could be modified to fit this setup was the Situational Awareness Global Assessment Technique.

8.2.1 *Situational Awareness Global Assessment Technique*

The Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1988; 1995; Salvendy, 2012, p.564) has been one of the most frequently applied SA measure with the capacity to assess all three of its levels. It has been used directly in experimental techniques that involve data gathering during ongoing task performance and is more commonly oriented around task suspension during an experiment (i.e. Choi, Ahn, & Seo, 2020). In task suspension all senses from the environmental input are paused, and the participant is probed into the status of all or select elements of the environment that were present at the time of suspension (Suss, Belling, & Ward, 2014). This has benefits over the SPAM as the user is not performing a scenario during the probing period. Operator-in-the-loop exercises are given and at the selected intervals for task suspension, the participant would then be given a battery of questions. Questions are generally delivered via a computer or by pen and paper. Data collected by the battery of questions are used to score participants responses, basing the correct or incorrect responses on the information provided by expert.

The SAGAT has been effective in domains including aviation (Endsley 1990a, 1990b, 1995), military training (Matthews, Pleban, Endsley, & Strater, 2000), and air traffic control (Endsley, Sollenberger, Nakata, & Stein, 2000). It has a predefined structure and is delivered in a systematic fashion. The SAGAT demands the identification of the exact entities and aspects of situational awareness to be assessed in the scenario, to then form a custom SAGAT application for that specific investigation. For example, the task-critical situational awareness cues are firstly needed. These cues comprise of the information in each environment that facilitate the perception and projection of future situational state(s). To extract these cues, a task analysis would be conducted, and the expected results would be the basis for the probe questions to assess participants' awareness of those cues. As the SAGAT is intrusive to task performance, there is possibility that stoppages influence the cognitive processes involved. Additionally, frequent stoppages would not occur in an operating room environment with similar reasons why stoppages in airplane flight tests are not possible. Consequently, the SAGAT may be adapted to be used during natural intervals during this project's experiment phase, along with being administered after the scenarios have finished.

8.3 Subjective Measures of Cognition-Based Assessment

Subjective measures have the ability to explore the *how* and *why* towards human performance, which may not be recorded by quantitative methods. A numerically orientated format can still be maintained, however. The most common metrics of mental state evaluation have been a subject's mental workload and task engagement (Guru et al., 2015). Performance-based metrics and self-report questionnaires have been the typical forms of gathering this information. Scales can comprise of one-dimensional workload measures such as the Overall Workload Scale (Hill et al., 1992) or the Modified Cooper Harper scale (Casali & Wierwille, 1983) (MCH). As such, the MCH is a 10-point unidimensional rating scale forming a global rating of workload (Hill et al., 1992. p.431). Alternatively, there can be multidimensional subscales pertaining to specific mental resources, such as the Subjective Workload Assessment Technique (SWAT) or the National Aeronautics and Space Administration – Task Load Index (NASA-TLX). These multidimensional examples of measuring workload are the two most commonly used subjective assessment means. As such, the NASA-TLX has specific mental resources subscales of physical demand, mental demand, temporal demand, performance, effort, and frustration. However, these measures must be performed at the end of a task as they cannot be administered in an unobtrusive manner during the same task. The reliability and accuracy of measurement may be decreased due to this constraint as biases can occur during subjective retrospective self-assessment within these forms of cognition-based assessment.

This outline of measures highlighted the problems with qualitative measures of cognition. Indeed, subjective measures may be more susceptible to biases. However, further considering this Hederich-Martínez, López-Vargas, and Camargo-Uribe (2016) created an online cognitive training application that provided users with feedback of their own cognitive styles. Learning styles, and a self-report for users' motivations were also measured and used as feedback. This feedback also held suggested actions that would be useful to each student in their academic studies. Delivered virtually, results for the participants receiving this feedback achieved better learning outcomes than the control group. Importantly the cognitive strategies were also observed to be modified slightly in the students with better cognitive understanding of their self-regulated learning strategies.

Therefore, although contradictory to reducing quantitative measures and increasing usage of qualitative measures in experimental testing, there are benefits for understanding effects on individual cognition. To bridge the gap between the mental models of a novice and an expert, trainees must verbalise their own reasoning and experts can then help modify their mental model to better match the experts. An effective application requires measures that can extract mental models represented within each scenario, that allows a learner to develop their understanding through clear explanation of thought. The method most suitable for this was posited to be Verbal Protocol Analysis (VPA; Ericsson & Simon, 1998; Ericsson, 2018; Korovin, Farrell, Hsu, White, & Ghader, 2018).

8.3.1 Verbal Protocol Analysis

Bainbridge and Sanderson (1990) described VPA as a method to capture the mental processes of human operators. It was intended to create valid references from the verbal content of a human operator (Weber, 1990; Ericsson & Simon, 1984). VPA has been used extensively within the Human Factors domain as a tool to gain insight into the cognitive underpinnings of complex behavior and meaningful inferences about interrelationships. Verbal protocol methods have been used to measure situational awareness in domains involving automated driving (Key, Morris, & Mansfield, 2016), nuclear power plants (Lee, Park, ryum Kim, & Seong, 2012), and emergency situational driving (Banks, Stanton, & Harvey, 2014). For example, Den Hartigh et al (2014) asked football players to watch a video of a football game and describe the events occurring that they identified as interesting. The football players had differing skill levels and through VPA they uncovered that more experienced players reported and explained the meaning of a substantial number of events that were away from the ball, with less skill players having a narrower range with the ball at the centre of their attention.

8.3.2 Concurrent Verbal Protocol Analysis and VR

Trickett and Trafton (2009) stated that VPA can be especially fruitful for virtual reality and other virtually simulated environments, and they suggested more use within this domain. The specific type of VPA most appropriate within this project was concurrent verbal protocol analysis. Concurrent verbal protocol analysis provides valid measures of an individual's metacognitive states during a task (Pressley & Afflerbach, 1995) and with use of VR does not interfere with the visual or tactile feedback required to interact with displayed tasks. This method uses a simultaneous output of a participant's responses to a task alongside their verbalizations of the task being performed. Concurrent methods in VR, with comparison of novice and expert usage, allows identification of real-time events to highlight areas where novices would benefit from modification or more support during the sessions. With the perspective that talking aloud represents the output of the contents of working memory, a researcher can also perform post-intervention analysis to identify details of the strategies, concepts and beliefs held, understanding of task goals and structures, elicited affective response, and many other cognitive processes. This is called retrospective VPA and another option towards capturing participant experiences.

Russo, Johnson, and Stephens (1989) investigated the effects of concurrent VPA with retrospective VPA in a gambling task, and an addition task. Retrospective VPA asks participants a set list of probing questions after their experimental session. They suggested retrospective VPA has weakness in that participants forget or fabricate cognitions and action in both tasks. This was a consideration again using retrospective VPA in the project, as although concurrent VPA was suggested to increase tasks time and alter performance accuracy, the importance of accurate capture of cognition outweighed deviations in performance measures. Caution must be taken when using concurrent VPA in that the instructions to participants are to talk about *what* they are doing, rather than *why* or *how* - this can interfere with their ability to perform the main tasks at hand, therefore effecting performance. When behaviour is affected by verbalizations in a single trail (known as first-order reactivity), measures to preserve the most important factor of interest should be put in place. In this project it is with the combination of post-intervention analysis/questions (to preserve cognitive processes during tasks), and caution in the quantitative performance metrics interpretation. The *how* and *why* should be found after intervention, without consequence to understanding the cognitive effects during the intervention.

8.4 Extracting Participants' Cognitions for Later Analysis

Lastly within this chapter, consideration is required for the method of extraction of participants' data. As participants work through the training material their unique data would need to be captured for analysis. As each participant may have a large amount of data due to both verbal recording and any quantitative measures, an efficient method of capture, extraction, and processing was essential to be feasible to analyse in the researcher's time frame. From a psychological research perspective, an ideal method would be to record the audio and video of participants to capture all information provided. This would involve a lengthy process of transcribing audio, pairing the video files with the related audio clips, then further analysis with the relevant evaluation tools/methods. This process would capture and maintain the most amount of data through the experiment and analysis and would be ideal if time permitting. Any material used in a subsequent control condition would also be captured in this process.

In this chapter, assessing the measures and deliberating other data related elements was intended to aid in designing the experimental materials as well as an experiments' design itself. The next chapter attempts to encompass the lessons learned up to the current point, create the final application to be used by operating room trainees, and create a simple but valid experiment and analysis process.

Chapter 9 - Design and Creation of the VR Training Material

9.0 Aim

The aim of chapter 8 was to create a novel NTS training application that could be examined for its ability to improve practical non-technical competences compared to current materials. As the resultant system was intended to be an educational program that enhanced the cognition of Operating Theatre staff the name given was Non-Technical Skills in Virtual Reality (NTS-VR), with similarity to the commonly used 'NOTSS' material but not specific to surgeons. The aim was to produce a preliminary application that provided support for the premise that both *explicit practical content and theoretically rooted information within it, can be taught to trainees*. This may be used to augment the currently deployed workshops and lecture-style courses across the UK.

9.1 Participants and Footage

For this to be created, an Operating Room recording of surgery was required along with its processing and post-editing to create the VR material. Therefore, two procedures were recorded at a hospital in Glasgow (for consent see Appendix Figure 11). The first was a Le Fort 1 osteotomy, and second was a Bilateral Sagittal Split Osteotomy (BSSO). The BSSO is an elaboration of the Le Fort 1 as it splits the Maxilla (upper jaw) into 2 or more sections, each adjusted differently. In total there was 4 hours footage and was encoded and kept on a password-protected computer. However, guided by the previous Task Analysis, the footage reduced the parts with more occurrences of NTS events. This reduced the length to 30 minutes footage; the rest was permanently deleted.

9.2 Ethical Considerations for Recording Surgery

The National Research Ethics Safety (hra-decisiontool, 2017) criteria defined research as 't*he attempt to derive generalizable new knowledge including studies that aim to generate hypotheses as well as studies that aim to test them*.'. For the aim of recording a surgery, no new knowledge from individuals in the NHS was gained, as only current knowledge was extracted and recorded (i.e. verbal information about the surgery). Therefore, any hypotheses within the project referred to the reuse of current knowledge, and how it can transfer this current knowledge. Upon completing the NHS and medical research council's online REC approval query tool (hra-decisiontools.org, 2018), this also stated that video recording and its usage would not be deemed as research therefore further NHS approval was not required. To confirm this, a clinical improvement officer for the NHS Greater Glasgow and Clyde area was contacted by the researcher and they advised the researcher with similar information. Finally, as a final check the clinical governance lead for NHS Greater Glasgow and Clyde was contacted and stated no clinical governance was stated necessary based on the project's details.

The University of Huddersfield HHS SREP website (hud.ac.uk, 2016) similarly stated that '*a project is defined as 'research' if it seeks to systematically generate knowledge to share with others including the academic community, external partners such as funding organisations and/or the wider public*'. The procedures and methodology chosen may systematically capture knowledge, but did not generate knowledge. This means no new knowledge would be produced from the content of video recording, but the educational content within the video was reused. Patient consent was gained with use of a standard NHS media/education consent form and one of the researcher's co-supervisor directly fulfilled this. The co-supervisor was a staff member in the location of recording. As further defined by University of Huddersfield HHS SREP website, '*It* [research] *may also include many evaluation studies, for example, evaluations of new services in health and/or social care, education and criminal justice*.'. For this project, there was no evaluation of any services, individuals or processes of the NHS on the part of the researcher. The information gathered was presented to academic students in an educational setting (also see 8.343 as VR content was edited). This information was related to their skills, and any participant being presented with this information will either already possess this information or will be attempting to possess it through their current surgical training course. No evaluation of the information or service occurred.

9.3 Apparatus and Materials

Investigating how to create VR footage was performed by Google searches and YouTube videos (Goolge.com, 2020; YouTube, 2020) along with reading articles that have created similar content. The researcher had no prior knowledge of any software however the final product was created with the subsequent setup. These steps have also been presented to be a guide for others to follow in best practice and to understand how to create VR training material in an operating theatre or similar healthcare setting.

9.3.1 *Computer Specification*

The software required high amounts of Random-Access Memory (RAM), graphical processing, and task processing by the Central Processing Unit (CPU). Therefore, a high specification computer was used with the following system specifications:

- Windows 10 Home Premium (64-bit) Operating System
- Intel® Core™ i7-8700 Hexa-core Processor with 3.2 GHz / 4.6 GHz and 12 MB cache
- 16 GB DDR4 RAM (2133 MHz)
- NVIDIA GeForce GTX 1080 graphics card (with 8 GB DDR4)
- 2 USB 3.0 ports with 2 USB 2.0 ports, DVI to HDMI 2.0, and HDMI 2.0 link
- 3TB HDD (7200 rpm)
- Software was installed onto an internal 250 GB SSD
- Corsair VS Series 450 W Power Supply Unit
- ShadowPlay screen/video/audio recorder software

Importantly, the ShadowPlay software was able to record the user's field of view in VR, so that their precise head movements and interactions were able to be recorded and analysed later. Any variation of these specifications will result in some form of loss of either quality or time, being that playback of an application may not have high Frames Per Second, or that media rendering time is longer and can potentially add days/weeks to media creation. ShadowPlay (or similar VR recording) has a similar optimal requirement to function.

9.3.2 Virtual Reality Headset and 360-Degree Camera

The virtual reality headset used was the latest commercial headset available at time of sourcing in Q4 2017. The HTC Vive (htc.com, 2019) has 1080x1200 pixels per eye (2160x1200 pixels combined) with a refresh rate of 90 hertz and 110 degrees Field of View. The Vive is paired with wireless handheld controllers which have multifunctional input buttons/triggers. With typical 2D video recording there is little quality loss when presenting on a 2D screen such as TV or PC monitor. However, when presenting the same 2D media in a 360-degree setup the media is stretched to fit the required space as a typical camera captures only a small field of view in degrees. This means that the video ratio becomes distorted and quality is significantly reduced if the media was played in a VR headset. To compensate for this, 360-degree video rigs use several cameras with equal field of view in degrees that combine to add up 360 degrees. Additionally, each camera should have a resolution that is equal to or above the resolution to the VR device the media will be presented on. At the time of selection, 2 cameras met these criteria: the Vuze (vuze, 2019) and the Insta360 Pro (insta360.com, 2019). The Insta360 Pro was chosen as it had 6 x F2.4 fisheye lenses- each have 4096 X 2048 resolution. This allowed a maximum post-processing stitching video resolution of 7680 X 3840 at 30 frames per second. The Insta360 Pro was better than the Vuze (with 3840 X 2160 per eye). To source funding for the rental of the required equipment, a proposal was submitted to the Centre for Applied Psychological Research (CAPR, 2019) at the University of Huddersfield. Upon review, the proposal was approved and the CAPR allowed the equipment hiring from a UK digital equipment rental company (Harmony LTD, 2019) for the length of the recording session.

9.3.3 Other Hardware

Due to the quality of each of the 6 videos, there were approximately 2 gigabyte per minute of data being recorded. Therefore, a high-speed storage device was needed. A typical Hard Disk Drive did not have adequate data transfer speeds to support the Insta360 Pro. Therefore, a high speed 500 GB SSD was used to record the videos as SSD's transfer data at much faster rates to HDDs. A monopod was used to hold the 360-degree camera instead of a tripod as its design was more slim line and would use less space in the OR therefore minimise spatial interference.

9.3.4 Software

An additional factor in the selection of the Insta360 Pro was the advantage in the video stitching process. After recording, each camera's 2D videos had to be stitched together with each camera at each side. To do this, stitching software typically overlays and merges the outer left 15 degrees and outer right 15 degree of each camera to result in the inability to identify any seams or lines where 2 videos would meet. The remaining 60 degrees of each camera becomes a 360 video (as there were 6 cameras in total). However, the Insta360 has 200-degree field of view per lens therefore could overlay more content which allows more accurate merging of each video. Upon rental of the equipment, the researcher completed a VR training session at Harmony LTD, which included the correct practice of the camera, SSD drive, and the Insta360 Stitching software.

Adobe Premiere Pro, Photoshop, and Media Encoder were the Adobe software used. As the researcher had no previous experience in such software, tutorials were sourced for free or purchased from several sources. Such sources were YouTube (youtube.com), Udemy (udemy.com, 2019), or books. Adobe also released an add-on whereby stitching the 6 cameras can be performed in Adobe, instead of separately in the Insta360 stitching software. This improved the quality as the videos are being encoded, decided, and codec wrapped less. Adobe Photoshop 2019 was used to create the materials within the surgical videos. Text, images, and other static media were created in Photoshop and saved within the project files. When these saves Photoshop files are imported into other Adobe software a dynamic link was created whereby the original media can be edited once, and any software using this media will also have the media updated. Therefore, Photoshop was used first to create all media, such as instructions, images, text etc. Indeed, the final video content was created by using a mix of video clips, images, and audio.

Premiere Pro 2019 supported editing in most formats such as 8k and Virtual Reality. It has been the industry standard software in video editing. To access Premiere Pro in additional to all Adobe apps a subscription was created via University affiliation. Although the computer used in this project was of high specification, certain preferences were needed in Adobe Premiere Pro. Proxy videos were used, which were low resolution copies of the original media -which is similar to halving the display quality of a video. Before exporting any completed media to Adobe Media Encoder, videos were rendered and previewed using GoPro VR Player which was an add-in playback option.

This material was presented in a non-chronological order and more importantly the context of the videos and images had been significantly distorted. This means that the information presented to the user did not convey the true events. For example, a user can be provided with context and material that a scrub nurse has not called out for a gauze pack to be recorded, and the video matches this information. The true event and recording may not have had this event occur and select clips would have been modified to depict this. The combination of custom-made audio paired with modified videos, which were different to the true chronology, enables a platform for flexible content creation.

Adobe Media Encoder 2019 was a transcoding application that allowed the researcher to output media to a chosen final format for distribution/usage. Once the project was exported from Premiere Pro to Media Encoder, through trial and error the researcher found the following settings were best for creating the final videos, for readers intending to create similar content: Equirectangular and monoscopic. 30 fps timebase, 5120x2560 (2:1) horizontal/vertical aspect ratio, a square pixel 1.0 aspect ratio, no fields (progressive scan), and a 30fps timecode display format. The audio was 2 channel stereo, and 48000Hz sample rate 3D audio was not used). Additionally, maximum bit depth and render quality were selected along with ‘composite in linear colour'.

As the HTC Vive VR display was a secondary display to a main PC screen, this increased the usage and need of the CPU and GPU. In addition, if a high-quality experience was desired then VR can demand intensive use of both hardware components as well as the RAM memory. To increase the efficiency of the decoding and rendering of VR videos, a relatively new codec was used that was not yet the standard in VR production in 2018 and 2019. H265, also known as High Efficiency Video Codec (HEVC) was a result of revision and improvement to the standard H264 codec. This H265 codec had less lines to decode, compared to its predecessor. From this, bit rate can be halved yet the same quality can be presented. However, use of this new codec required a high-performance computer that was able to support the 'backend' framework needed (Malhotra, Singh, & Matam, 2019).

Once all the media were created and exported, a delivery method was needed with a menu driven system and interactive elements able to be built into the program. Two predominant software applications used are Unreal Engine (unrealengine.com, 2020), and Unity (unity.com, 2020). Both engines have been used for world leading games and professional applications (artinstitutes.edu, 2019). Simplifying their functionality, Unreal Engine has stronger comparative features to 3D computer graphics and gaming with ultra-realism when compared to Unity. The primary benefit for this project was the use of 'Blueprints'- they are the Graphical User Interface that converts coding in C# to a user-friendly interactive system. This was advantageous for the primary researcher who had beginner level skills in such programming language. Unreal has a user-friendly design that would be able to support the simple needs of their current project in a short time period. More elaborate programmes should use Unity engine as Unreal has limitations that will significantly prevent creation of complex cognition-based apps. Unreal engine 2018 was used to create the application.

9.4 Final Experimental Material

The final material allowed participants to play and pause the VR material, with recording of their session. Participants were able to see most of the detail in the environment however specific writing was not legible. Users would identify key cues they believe were important to NTS and patient safety (see Figure 12). They would also verbalize the reasons for each cue selection. After the video was completed, a summary feedback from a more experienced individual would then be shown to compare reasons for choices. Below are the designs of the material (Figure 13 and 14), however the final material did not have the menus as presented as this allowed the participants to freely categorize their selected cues without reduction to a limited amount of options. The material was uploaded to the University of Huddersfield student OneDrive storage and can be viewed upon request.

Figure 12

The initial designed layout of the training material- there were not additional text or images until the user interacts with the video.

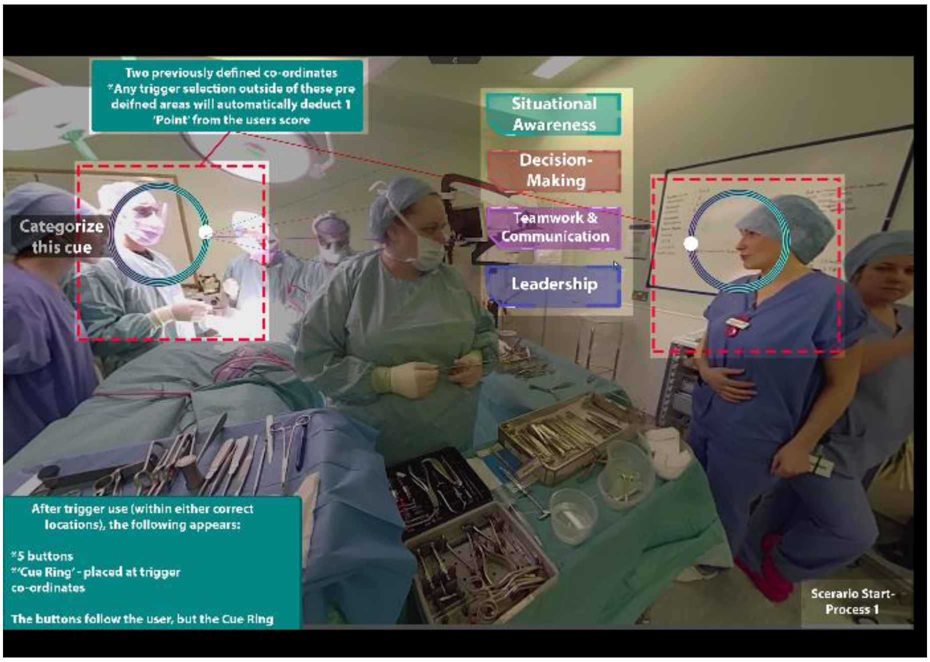


Figure 13

The Unreal Engine Blueprints created to run the VR Application. These were the links between VR setup and the 4k 360 degrees OT videos, along with the interactive functions.

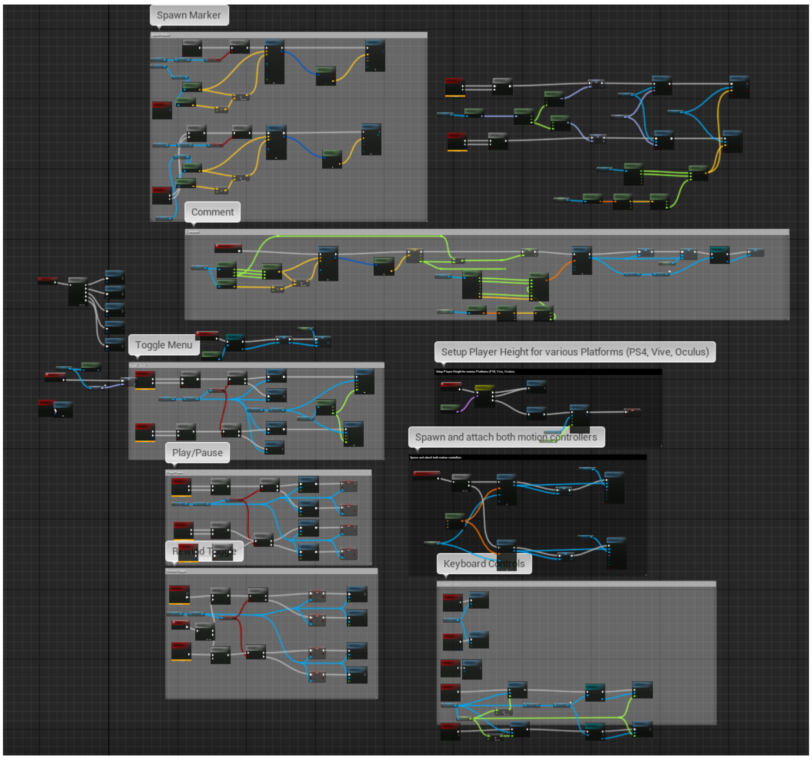
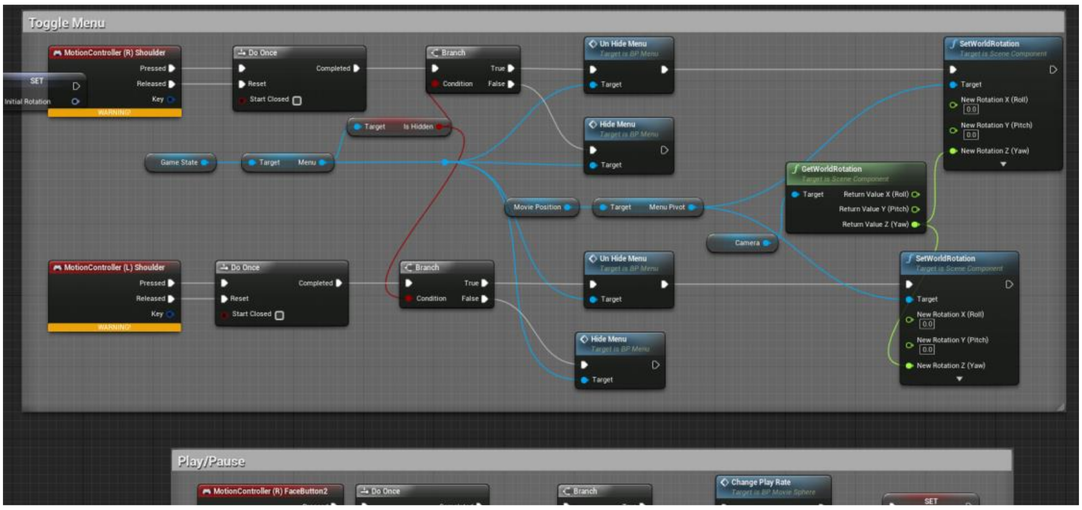


Figure 14

A section of the application’s structure for the menu system- taken from Figure 11.



Chapter 10 - Experiment Information

10.0 Aims and hypotheses

Although the application was created based on the theoretical and practical information retrieved and analysed in previous chapters, there was still a requirement to explore whether the application had benefits over current methods. There can be risk in producing new tools that were not founded upon empirical and/or exploratory results and the misappropriation of usage may occur due to this lack of insight. Therefore, this chapter described the investigation involving 2 groups of undergraduate operating room trainees either in a control or experimental condition. From reviewing the literature into the applications and theories of performance enhancement in simulation, the main aim was to assess if there was a change in both perceptual and cognitive ability for participants based upon the condition. These changes were measured by the comparison to expert-driven information, along with evidence from Human Factors-based data analysis technique to identify any change in cognition.

From the coding of the cognitive verbalizations there was the opportunity to explore if the frequency of NTS related cognitions increased in the post- condition measure compared to the pre-condition measure for both groups. Therefore, a hypothesis was formed based on the notion that participants in the control condition would not apply concepts/theory to real world application/simulation as efficiently as participants in the experimental condition. It was hypothesised that *participants in the experimental condition will have a significantly increased number of NTS related cognitions compared to the control condition, as measured by the encodings from participants' cognitive verbalizations.*

On a qualitative level, it was also predicted that the participants in the experimental condition would produce improved meanings for safety related cues in their cognitive processes compared to the participants in the control condition. This was because the technique to which participants were taught would better scaffold their ability to comprehend the properties of cues/events compared to the control condition.

10.1 Participants

Fourteen undergraduate students were recruited via email to the Operating Department Practitioner course leader and subsequent student emails. Thirteen females and one male had a mean age of 22 years, with a range between 20-25 years. The ages of participants were 18+, with an approximate age range of 21-23 for most participants. Seven participants were in each condition and all participants were in their 3rd year of Operating Department Practitioner undergraduate course at the University of Huddersfield. All participants have assisted surgeons in more than 50 surgeries in an array of procedure types, as the course required them to assist in several surgeries per day.

There were criteria for participants to be from the same course year, as this should reduce the variation in non-technical skills. They were initially asked for their knowledge of surgical instruments, the procedures being taught, and other experience and knowledge-based information. This was to approximate the individual's abilities in the *novice* category in reference to the Dreyfus model (Dreyfus & Dreyfus, 1980) as previously described. These were evidenced through the results of the previous Task Analysis. For example, they needed to know basic surgical tools, and basic operating theatre protocol to ensure that upon allocation of condition, all participants had parallel knowledge. This aided to reduce confounding factors that may have affected the data if disparity was present. Additionally, if participants were over/under skilled the potential benefits of the simulator may not have been captured- as it would aim at a population with comparatively sub-optimal effectiveness. Most of participants stated they could view most of the procedure in the OR, with four participants stating they can see all the procedure. This implied that they had opportunity for planning ahead when helping the surgeon as they knew what tasks were occurring. These views were equally split between the groups. However, participant 10 stated they could view few parts of the procedure and this was considered in the discussion for the results of participant 10 as their predictive abilities *may* have been less developed. Although in their 3rd year and should have been homogenous in technical skill levels, the researcher performed this check to find if any individual did not have a corresponding level of understanding of the surgery shown in the experiment. All participants had assisted in surgery and stated they were competent with theatre procedures and equipment.

All participants stated they had satisfactory vision to read text within the surgical simulation. To be eligible, participants also must not have had suffered from epilepsy or Attention Deficit Hyperactive Disorder or had been on any anti-psychotic or anti-depressant medication. Participant were made aware of this and verbally stated they were able to continue. Participant's age, sex, and video/audio recording of their sessions in the experiment were recorded. Results of questionnaires exploring previous surgical experience, technical, and non-technical knowledge were also recorded.

10.1.1 Justifications for Concurrent VPA with Participants

As concurrent Verbal Protocol Analysis was used in this project (see 7.5.1) Trickett and Trafton (2007) stated that VPA typically includes fewer subjects than other kinds of psychological studies. Contributory factors typically have been that experts can be scarce and difficult to recruit, or that due to the richness of data it can be voluminous therefore analysis may be restricted. Over the course of the development and implementation of VPA, approximate guidelines have suggested 1 expert and 1-3 novice participants be included if forming hypotheses. With the absence of experts this number should increase. Trickett and Trafton (2007) suggested 5-10 novices can be an ideal amount for investigation, and with more participants over this number posits a substantial amount that may be better analyzed with means of a coding scheme and a subset of the data to be coded and analyzed.

As this thesis reviewed previous research that established initial theories of non-technical skills, there was requirement for more participants to facilitate richer qualitative and evaluative quantitative analyses. Therefore, this project selected to code and analyze 14 novice participants which was a feasible amount of analysis for the timeframe allowed. There was also data collected from the 3 previously mentioned CAPs, therefore utilizing a bigger pool of expert-based information in this project than suggested. This number of participants was 2 less than Korovin, Farrell, Hsu, White, and Ghaderi (2018) who also used concurrent Verbal Protocol Analysis to examine the differences in thought processes between experienced and novice surgeons when presented with a critical situation. Lastly, one individual was an experienced general surgeon at Leeds General Infirmary with several years of teaching in surgery and had been previously on many non-technical skills courses. Consequently, this individual held both theoretical and practical competence, which was suitable to be used as feedback. She was able to meet the researcher and provide her cognitive processes to be feedback material for participants in the experimental condition. The feedback was added into the VR application created in chapter 8. Importantly, there was no direct mention or explanation of specific theory but were embedded in the context of her verbalizations.

10.2 Design

There were 2 conditions – an experimental condition and a control condition. The conditions were exposed to only a single set of experimental material therefore was an independent samples design. For the quantitative elements being explored the independent variables were the type of material used (control/experimental), and the dependant variables were the SAGAT scores and coding outputs from the cognitive verbalizations. The sampling method was opportunity (aka convenience) sampling but it was also snowball sample as two participants informed their colleagues of the study thus participants were recruited via word of mouth. As the participants were briefly assessed for a basal understanding of key elements on the OR, along with approximation of their decision-making abilities the sampling attempted to increase homogeneity. Based upon the session suggestion for optimal training per session by ShadowBox (Klein et al, 2017), this project identified the optimal delivery frequency, length, and intensity was suggested to be short (one or two scenarios per session), distributed, and over extended time periods. However, due to time and resource limitations, along with being exploratory more than real-world training, as well as doubling the amount of data analysis required, only one single-session was performed. The consequences of this are expanded in the discussion section.

10.3 Apparatus and Materials

10.3.1 *Information Sheet*

The information sheet was provided to most participants several days prior, however some participants were opportunistically recruited with the information provided the same day. It presented an outline of the project however avoided the keywords of 'non-technical skills' to reduce the chance of participants investigating and learning the key theories and concepts beforehand. The terminology surrounded their general awareness of events when in the theatre, and their feedback on virtual reality usage. However, to ensure accurate disclosure before engaging with the experiment, participants were verbally informed of the specific nature of the study in their session including the interest in their non-technical knowledge and abilities.

10.3.2 *VR Equipment and Participants Demographic Information*

A HTC Vive along with the relevant PC hardware (HTC.com, 2016) and software (Unrealengine, 2020) was used (see section 8.4). Using ShadowPlay software (NVidia, 2018**),** the user's views and actions during their sessions were recorded and stored onto the PC. The Vive could be modified for items that may have hindered optimal usage, such as glasses or religious headwear etc.

General participant information was recorded to ensure there was not the presence of a participant with considerably more or less experience both in theatre and with VR usage. The information was as followed: age, sex, stage of study, surgical specialty, amount of procedures they have assisted a surgeon with, amount of procedures they have observer in surgery, how well they can view events in surgery, where they find their sources for learning more about surgery, what technology they use in education, the amount of time this technology is used, experience of using a VR headset, and confidence about key elements in surgery.

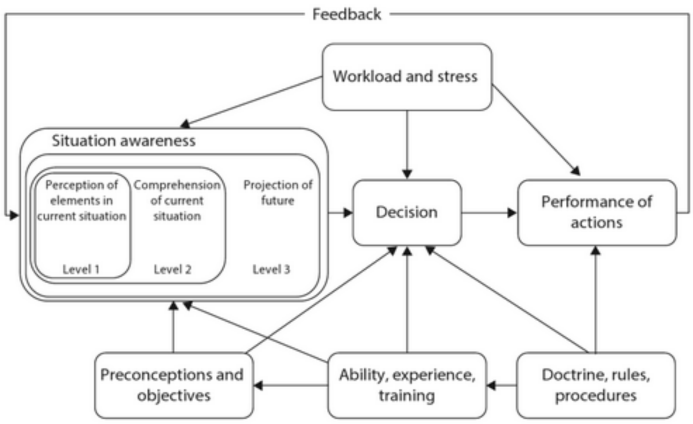
10.3.3 *Decision-Making Scenario Ranking Test*

Decision-Making has been described to be the output of a variety of internal and external mechanisms which influence the cognitive processes of the individual (Endsley, 1995). Such factors as the goals, environment, and the individual's understanding of a system can effect decision-making. From this, if the same scenario was provided to individuals with an array of skill levels, each participant response made could be analyzed to understand such pieces of information used to form their final decision. At its basal form and coinciding with the 4 groupings of decision-making (as discussed in section 3.4), decision-making can be captured in 4 categories. For this project, the global indicator of participants' strategies used was the explained decision from their proficiency to obtain information, understand the dynamic links between each piece of information, and select a goal-driven option. Therefore, following from the written scenario provided to the previous 85 participants in the previous data collection, a decision-making scenario was created and reviewed by the three CAPs. As mentioned in chapter 6, these CAPs were interviewed to gain insight into their cognitive abilities and practices. Prior to the start of their cognitive interviews, the CAPs were provided with the same decision-making scenario on paper and asked to rank the order. They then typed out their rationale for each rank selection as the researcher prepared for their cognitive interview. To ensure the researcher had accurately created 3 solutions that each matched their desired type of decision-making strategy, the CAPs were also asked to state which option best fit into which strategy. They individually concurred that each option had the intended strategy embedded within it, however option 3 was modified by adding the phrase '*you know that this is standard practice in this situation',* to improve its justification for its more urgent approach.

The CAPs ranking order and explanations could then be used as comparative data, to understand what differences occur, why they occur, and how they can approximately gauge non-technical strategies used per participant. Once completed, the decision-making scenario answers each represented either a rule-based, recognition-primed, or analytical underpinning. A creativity-based answer was not provided as due to its terminology, its solution could have been used as an 'opt out' if a participant was unsure of a choice. The prediction was that all participants should have homogenous cognitive abilities in this project for each node in Endsley's (2000) situational awareness theory (see Figure 15) if their non-technical skills were equal- as approximately represented by the type of decision-making strategy produced. This was performed before any intervention material was shown, to approximately assess the strategies each participant held.

Figure 15

Emphasizing the position of decision-making, the model by Endsley (2000) suggested it is a product of several factors; each factor assembled/understood differently by each person therefore differing decisional strategies and output will occur.



10.3.4 *Verbal Protocol Capture*

Concurrent VPA was used and is discussed in 7.6.1, and 8.3.1. VPA analysis typically starts with participants performing a scenario and verbalizing their cognitions as they work through the scenario. This matched the setup of the current experiment as participants were instructed to talk aloud about their cognitions as they worked through material in each stage of their session. For the entire session, no feedback was provided to participants on their ability to follow verbal protocol. If they stopped talking, the researcher did not prompt them in any way. In both conditions, the participant was able to pause the material presented to them and discussion the content at that timeframe but also talk without pausing the videos. Based on Ericsson and Simon (1993) (see Ericsson & Moxley, 2018 pp 286-301 for a review), this process was used as the participants spontaneous attention to information with immediate verbalization of their thoughts was expected to be the closest connection to their actual thoughts before talking. The scenarios presented to participants were cognitively complex and fast paced enough that they likely would not have time to deliberate on what they have said, are saying, or will say next. This meant that participants verbalized their thoughts as a representative stream of information that occurred at that moment in time. Park (2010) established that VPA depends upon a well-defined speech act coding scheme to better distinguish the nature of the embedded data. This was also fitting with the current project as the verbal information was matched with millisecond precision to the actions in the simulation along with a categorized list of non-technical events that could form the coding scheme.

An example of coding in VPA when identifying underpinning theoretical components would be if a participant stated, '*it looks like she might get called away in 5 minutes*'. This would translate to the category of situational awareness, and the more specific element of Level 3 situational awareness being *prediction* as the participant had the capacity to understand the environmental contexts effectively to anticipate future events even though they may not be obvious. In the control condition, participants watching the video were able to pause and reflect on the point of theoretical explanation provided. For example, they may state they did not know that decision-making has 4 categorized forms of strategy and expand upon the usefulness of their new knowledge. In the experimental condition participants were able to pause the feedback of the video and also reflect on the elements provided.

10.3.5 *Non-Technical Skills Global Walkthrough Measurement*

To achieve a representation of each participant's non-technical understandings including their perceptual abilities, all participants were instructed to interact with a 360-degree video and identify anything in the video they considered were related to non-technical skills and patient safety. If participants asked further into what non-technical skills meant (as they may have wanted clarity on what events they were looking for) the researcher stated they should '*identify anything that you think is related to the safety of the patient and error prevention that it not directly the 'hands-on' actions of the surgeons'*. The video was 5 minutes long and the computer screen and audio of participants were recorded. The participants paused the video and used their mouse cursor to highlight the area of interest and provide explanation of this choice. This can be termed an intermittent cognitive walkthrough. The pertinent information that was presented from this measure were the verbal explanations provided for each cue- these were transcribed and analyzed using Verbal Protocol Analysis.

10.3.6 *Situational Awareness Global Assessment Probe*

A probing list of 28 questions was created that consisted of many different cues in the environment found in the video. The questions were based on the SAGAT technique and were modified to be relevant to the subject area of the project. The cues were divided into the different non-technical skills (approximately 6 for each category) and were selected based upon their relation to the events in the video. For example, a teamwork/communication cue question was '*at any time, did the surgeon ask the runner for help*?'. A situational awareness cue at level 1 was '*how many swabs were on the table near the end of the video?'.* Along with previous statement to not look for anything specific as the questions may be different. There were 4 questions that were designed to divert the assumptions of the participant, two questions consisted of fictional items, this was to assert if participants showed any researcher bias in that they would guess or infer the answer instead of legitimately identifying it. One question was '*what color was the poster near the doors*?', however there was no poster in the video. Two other questions were distractor questions. The previously mentioned method of observer-rating freeze probe recall, also known as 'real time probing', was not used as this would have intruded on participants performance, instead the probing list was administered by interruption of the video at the 2 minute and 41 second mark. Participants were verbally asked each question quickly after the scenario froze and computer screen went blank. This was not expected by the participants.

10.3.7 *Non-technical Skills Handbook and Human Factors-Based Theory Video Content*

For the control condition, a handbook was created for participants that described each of the non-technical skill and broke them down into further explanation. This was to be representative of the current content and medium used in such healthcare courses. The handbook was then accompanied by video content. Both the handbook and video content were customized so that the necessary theory was present which similarly paired to the NTS events in the experimental condition.

For the control condition, cognitive and social explanation of non-technical skills events were included in the material. The theoretical information was generic regardless of the institution/company or other source delivering the information. In other words, it did not matter if the material used was a recording of a lecture, a workshop, a 2d illustration, or other forms of media. The key point was that each theory was being accurately and concisely described so the concept was transferred to participants. For example, Endsley's (1995) theory of 3 levels of Situational Awareness, or how distractions can interrupt working memory. The underpinning theoretical information was required that equalled the same information that was embedded in the feedback in the experimental condition. This also had to represent the type of material available at the time. Most of the publicly sourced content (i.e. YouTube and Google) surrounded training material derived from aviation training. However, there were healthcare specific consent publicly available online from a leading NTS training institution that the researcher was able to use under the Fair Use Policy on YouTube (Youtube.com, 2019). Therefore, several links were saved to the relevant time ranges of each specific video which described necessary theory for the control condition (see Appendix). This resulted in a 20-minute compilation of a range of videos which represented each main concept in non-technical skills research.

In the experimental condition, an experienced surgeon had previously watched the same videos as the participants and their feedback was the comparative information from the material in the control condition. The format for delivery of non-technical skills related information was one way that differed from current training materials in use, represented by the control condition. Instead of being explicitly separated into descriptions and delivered by a lecture, book, or workshop, the utilisation of an experienced individual describing their cognitive processes allowed inference of the theoretical underpinning by example of real-world context. An example of the feedback was, *'the scrub nurse is already doing 3 things, but she will forget to complete the first task because she is about to get interrupted. She can only remember so much when it is busy like this.'* This also included the principles behind level 3 situational awareness, and how distractions can interrupt working memory. The combination of practical illustration but embedded cognitive representations was a novel method and its effect compared to the control condition were of interest.

10.3.8 *Non-Technical Skills Global Measurement*

With one of the project's goal to explore how both the experimental and control conditions affected participants' perception and comprehension, there was need to recapture their perceptions and cognitive processes. In this case, there was interest in exploring the change of understanding directly due to the content and delivery method in each condition. Therefore, another video was presented to all participants with instruction to identify anything in the video that they considered were related to non-technical skills and patient safety. However, the instructions dictated that the video may or may not be longer/shorter, the video may or may not have been edited (to change certain tools/events), or that questions after the video may or may not be different. The cues in the previous SAGAT questionnaire could be meaningless in the post-intervention SAGAT therefore participants were instructed to not consider this a factor in their performance during the scenario.

10.3.9 *Interpretative Phenomenological Analysis of Reflections*

The participants were asked to reflect upon their experiences of the intervention material in any way they would like. Their responses represented the key phenomena they believed were of value to include in their reflection. The intended type of exploration guided the choice of research method to be phenomenological in nature. The aim was to understand how the first-person setup of the intervention material affected participants. An idiographic methodology was therefore also required. One such method was Interpretative Phenomenological Analysis (IPA; Smith et al, 2009). The idiographic nature of IPA permits subjective and idiosyncratic-based exploration (Smith & Eatough, 2006). The participants reflections include relevant objects, events, and relationships such as surgical staff communicating, passing tools, and teacher trainee dynamics. They aid to give meaning in the lifeworld of the individual, with their ontological positions taken upon experience. This allows better understanding of how this intervention effects participants compared to quantitative outputs alone (Smith, 2004; Todres, 2007; Smith et al, 2009).

IPA is a contemporary qualitative analytical method methodology based on philosophies by Gadamer (1989) and Heidegger (1996). Hermeneutics, ideography, and phenomenology (Larkin et al, 2011) are its three underpinning foundations. Phenomenology determines the embedded perceptions of an individual, due to the multi-perceived nature they view the relevance and importance of their environment. Their reflections are their embodied and situated nature as they relay encountered experiential significance. IPA can provide insight into these embodied events positioned within participants (Larkin et al, 2011). These phenomena can be connected to psychological theories that can educate the effects of the intervention material- creating deeper exploration in terms of the fundamental experiential realities communicated, across global research outputs.

Upon analysis of transcripts, the use of analytical techniques such as abstraction, subsumption and numeration permit experiential realities to be made conscious, and mapped into hierarchical themes (Smith et al, 2009). Techniques which can be found in both VR and human factors theory aided in extracting experiential themes. Such themes were the sense of passing time (temporality), spatiality, and the subjective corporality (bodily) experiences during use (Todres, 2007). These factors were theoretical drivers to help the researcher understand their senses upon reflection, and allow the results be ideographic, transparent, and coherent when reporting the results (Yardley, 2008). Found across all experimental condition participants were XXX second level themes (See *Table 1).*

10.4 Procedure

Before meeting participants, equipment was setup to either be the material for the control condition, or the setup for the experimental condition. Participants were scheduled individual sessions and met on the University of Huddersfield campus. They were informed of the experiment and given the consent for to sign. Any questions were answered before signing the consent form. Questionnaires were provided and participants completed them with occasional question. The question which occurred for every participant was the clarification that a ‘head mounted display’ referred to the VR headsets. After this, the premade set of instructions were read aloud which described what the participant had to do during the pre-condition measurements. All participants received the same instructions.

Participants were allocated on their condition based on sequential participation- the first 7 were in the experimental condition and the remaining were in the control condition. Their existing abilities were measured with the decision-making questionnaire and the pre-intervention video before any intervention material were delivered. Participants were not informed they would be asked some questions about the pre-intervention video after they had worked through it. However, immediately after completing it they were asked the '*SAGAT Cue and Event Probing Questions'* list. At this point, the researcher then stated the participant would be in 'group 1' or 'group 2'- these were used to reduced researcher bias from using labelled groups i.e. 'control group', as this may cause participants to infer the condition may/may not be the 'better' one and form a bias predisposition of effects.

The researcher then set the pc or VR screen to record the user's session and stated the participant can press the spacebar to start and stop the video when ready. The researcher sat away from participant at another pc with a partition between them and was not called over until they finished. For experimental purposes to verify that all parts were exposed to the participants, they were asked to work through material in a systematic manner as to not skip any parts and were also not allowed to rewind. Rewinding would permit a secondary exposure to both audio and video for both conditions and the impact of this may be weighted differently between the conditions therefore may have confounded data if one group was to excessively rewind the material.

10.5 Analysis

The data were captured by synchronising the recorded screens with each user's verbalizations. These were matched up post-experiment to get a millisecond precise output video of what they were thinking/viewing, when each event occurred, each user's description of what they were thinking/doing, and why. A University of Huddersfield dual staff member and PhD student in the Department of English, Linguistics, and History provided transcriptions of the audio files. This individual provided copywriting and transcription services that would increase the accuracy of the audio to text transcription compared to the lead researcher's attempts. For Interpretative Phenomenological Analysis NVivo 20 was used to group participant quotes and form categories.

10.5.1 Verbal Protocol Analysis- Size and Structure

As explored in 7.6.1, concurrent VPA was used in data collection. The exploration of these verbalizations captured from concurrent VPA provided a synergy with the quantitative data recorded. This allowed a finer-grained analysis of both perspective and performance. Combining VPA with the quantitative results from each user allowed understanding of *what* was of interest to participants, but also *how* this affects their non-technical abilities. After transcribing the verbal outputs into a written format, the next step in concurrent VPA was to decide the size of 'grain' to use, to categorise the content. These may be the individual words, word senses, phrases, sentences, or themes (Chi, 1997).

The larger size to encode was at 'theme' level. This meant to encode the meaning of phrases and sentences into short thematic segments (Weber, 1990). For example, '*The scrub nurse counted 3 out loud, and the runner supervised this counting'* could have the underpinning meaning of addressing the *teamwork-based action* of staff. The thematic segment may be called *team-based safety behaviour* to *update status.* Stanton, Hedge, Brookhuis, Salas, and Hendrick (2004) stated that a VPA based on themes provide the most flexible and richest data source. However, there could have been benefits from the further reduction of grain size by further breakdown of a sentence (Pagliuca, Monaghan, & McIntosh, 2008). Indeed, verbal protocols have been analysed by a relatively new cognition-based Discourse Analysis (Cognitive Discourse Analysis; Tenbrink, 2014) that uses the full scope of methods for each grain size. As the analysis of sentences can explore general themes, sentence size was the optimal size of grain used and helped to keep within the time constraints of the project. However, a range of grain sizes were used where possible to better explore the participants' conceptual perspective (Schober, 1998), as they may not be aware of the effect of their language on the cognitive strategies used, when broken down further. Syntax (the arrangement of words) was not investigated in this study instead the semantics (concerned with meaning) were of primary focus.

10.5.2 Theory Based Coding Scheme for Analysis

Determined by the research question, the themes were encoded with rationale surrounding the theory and approach in chapter 4. This meant the encoding scheme was based on the NTS subject area including the 3 levels of situational awareness, the differing types of decision-making, teamwork and communication strategies, and leadership behaviours (i.e., See Figure 16). This helped to maintain construct validity, whereby this scheme setup was measuring verbalizations based upon an empirical, justifiable, and a contextually appropriate approach. An example of the same format was used by Rose, Bearman, Naweed, and Dorrian (2019, p16). For Situational Awareness, data were coded based on the three levels of Situational Awareness by Endsley (1988), with an added '4th level' based on Cristancho et al (2016). Perception, comprehension, and projection of future states were these three levels with *transformation/reflection* as the 4th level. Words/phrases/comment were categorized into these three levels before being further delineated and specified. For example, a verbal reference to a visual cue (i.e. *'oh in fact the monitor number is 80'*) was coded as Level 1 Situational Awareness. Understanding of the dynamics and actions within the environment were coded at level 2. For example, understanding standard communication procedures (i.e. '*What's she writing on the board? She hasn't told the scrub nurse about that.'*). Lastly, to project their understanding of the environment to be able to predict an upcoming action/event was coded as level 3 (i.e. '*if she doesn't count along I think they will miss that extra swab'*). Upon practice of the first version of a coding scheme, it then was developed and iteratively refined with feedback from the second rate (inter-rater) which modified some codes and formed merging of other codes to obtain a final scheme (see Figure 17).

10.6 Inter-rater Reliability

10.6.1 Importance of Measures

Once this coding scheme process was completed, the reliability of the coding was established, which was performed by assessment of reproducibility. This form of reliability can be defined as the consistency of a measure in that multiple independent assessments of the same event can produce a certain level of approximate scoring with each other (Cook & Campbell, 1979). Inter-rater reliability (IRR) between the 2 raters in this project was important due to usage of themes, as they relied on shared meaning. If the coding scheme was not well-structured, this could have caused drifting in meaning and an analysis of reliability would highlight inconsistencies within raters. The difficulty of capturing the variable skills in NTS for each healthcare individual placed greater importance to IRR as it represented the principle that one researcher's analysis of cognitive verbalizations should also lend itself to be re-captured by another similarly trained individual to a consistent level. This meant an independent rater was theorized to have strong correlative results to the same encoded data by a separate independent rater in a blind conditional setup. Before any encoding was started both independent raters made use of the in-depth instructions created by the 1st rater. This also helped to maintain good intra-rater reliability, so to keep consistency in rating by each rater frequently referring to these in-depth instructions to uncover any unintended refinement in analysis. Typically, raters compare their first and last sets of data analysed.

10.6.2 Randomly Selected Participants

It was suggested that the amount of double coding for IRR can be determined by the nature of the coding scheme and participants' transcripts. More double coding may be necessary if the codes do not occur frequently in the data. Trickett and Trafton (2009, p343) suggested 20-25% of the data is sufficient in the content of both Verbal Protocol Analysis and other similar techniques. This translated to 4 participants for this thesis, however as there were variability in the number of words per participant's transcripts 5 participants were randomly selected which would counter any issues if 2 participants with the minimal data were selected. Therefore, a random number generator (numbergenerator, 2020) with the range 1-14 was created and set to return 6 values. One value was used in the practice for the two raters before the intended 5 were analyzed. The values, hence participant numbers, were 3, 6, 8, 9, 11, 12.

10.6.3 IRR Requirements and Cohen's Kappa

At the highest tier, the data was set to be a 2 (2 raters) X 5 (5 NTS categories). This was performed for 5 participants. After the indices of measurement were set, the raters attempted to have equivocal understanding of the coding scheme and informally assessed participant 9- which was not included in the final inter-rater checks. Coding was refined until there was suggestion of an adequate and equal agreement level. The lead researcher changed more coding than the second researcher which iteratively instigated the lead researcher to recheck the other data carefully. Once both raters were in adequate agreement, they independently analyzed the 5 selected participants' transcripts. The results of encoding of the same material by two separate encoders was then analyzed using Cohen's Kappa correlational analysis. Coefficients of .90 or greater are strong in the majority of research contexts, with .80 being also greatly accepted in most situations (Trickett & Trafton, 2009). However, for exploratory studies there may be an increased variation due to the novelty of data which may cause deviation from the subjectively derived judgements shared across the 2 coders. This could result in a Kappa around .70 and could be justified for the exploratory indices in select studies. The assumptions to be met in order to used Cohen's Kappa were that the responses by each rater were measured in a nominal scale and the categories needed to be mutually exclusive. Secondly, the response data can be paired observations of the same phenomenon (the same instructions/data were given to both raters). Third, only two raters were involved and they analysed data independently. Cohen's Kappa correlational analysis results were presented in chapter 10.

Figure 16

Coding strategy breakdown of surgeons' and nurses' verbalized data used by Mitchell et al (2011). This thesis produced a similar coding scheme with the 3 levels of coding depth.

	2	3	4
Cognitive	Situation awareness	Listening	N You might pick that up by listening to the anaesthetist telling his helper
			S I think a scrub nurse who isn't listening and is very focused on their own tasks or worse still not focused at all on anything is a real problem
		Watching	N Constantly I'm looking at where the swabs are
			S if they're not watching and not paying attention they'll be handing you the scissors when you don't need them
		Understanding	N If you are following the operation and you understand the anatomy then you know [what stage] they [surgeons] are at
			S even if you deviate slightly because there's a slight anatomical variation in the patient she will realise because you have started to do that … you are re-doing step Y so therefore you need the instruments for step Y
		Anticipating	N I don't like the surgeon to have to wait for things. I like to have it there before they need it so things run smoothly
			S You just put your hand out and you've not even opened your mouth to ask for it and what you need is there.
	Decision making	whether to act or not	N Deciding what equipment to take on [the trolley] at what point
			S No data coded
Social	Communication	Verbal	N I would say if I'd lost a needle in my counts or if he hadn't given me back a needle
			S I probably would have liked to have known in advance of positioning something perfectly before they say, "actually I can't find instrument X"
		Non-verbal	N Sometimes I'll pick up a fresh swab to indicate [to the floor nurse] that I need more
			S When I say communication, you know sometimes there isn't much conversation at all, it can be virtually silent
	Teamwork	Sharing	N You have to keep the floor staff in tune with what is going on at the table
			S I could have easily lost the patient on the day but thankfully we didn't and I think … not only because of me … the nurse and the anaesthetic colleagues worked as a team and we pulled the patient out of deep trouble.
		Clarifying	N I might say [to the surgeon], "that's not how we normally do it" but you always try to be tactful
			S Ask questions … say that you don't understand it because most [surgeons] would be quite happy [to explain], rather than trying to bluff your way through something,
		Co-ordinating	N I always think if I have a good floor nurse, things will go smoothly
			S I need a scrub nurse that's thinking with me, preferably
	Leadership	Guiding	N If it's a more junior surgeon, they might look to you to ask what does Mr X [consultant] use here?
			S it's the theatre team that should have that [problem] resolved but the scrub nurse, I guess

Figure 17

The coding table used in this project which captured the non-technical skills of participant. This was based on the NOTSS taxonomy, the structure based upon Mitchel el al (2011), and Korvin et al (2018).

Overarching Categories	Generic NTS	Specific NTS Element	Cognitive Ability	Code	Example Recognisable Occurrence	Count
Cognitive Processes	**Situational Awareness**	Perception	Perceiving Human Action	CSPE1	States actions performed by staff	
			Perceiving equipment processes	CSPE2	States equipment usage or status	
		Comprehension	Identifying embedded dynamic interactions	CSC1	Stating the dynamics of 2 or more elements	
			Process Deviation recognition	CSC2	Describes what 'should occur'	
			Other individuals Increasing or reducing SA	CSC3	Identifying others disconnecting from primary tasks	
		Prediction	Deliberation- practical or theoretical	CSPR1	Talking aloud of what could occur next	
			Detailing and justifying a prediction	CSPR2	Giving more detail to the predicted events	
		Transformation	Cognitive advancement in event/cue meanings	CR1	Gaining more info from same cue, due to processing change	
	Decision Making	Information gathering	Identifying cues or information towards a decision	CDI1	Acknowledging visual or audio they state have meaning	
		Understanding Options	Describing the consequences of a decision	CDU1	Explaining the new line of events/consequence, due to a choice	
		Selecting a type of decisional logic	Rule-based decisional behaviour	CDS1	Next task identification. If/then logic.	
		Tacit assessment	Creative/ Analytic logical assessment	CDS2	Conceptual or tacit event logic	
		Indecision	Inability to understand and choose best option	CDI2	Verbalises doubt/unsure. DM strategy not right	
Social Processes	**Teamwork**	Environment interaction with 2 or more	Patient safety related, and content specific events	STA1	Explaining how the teamwork affected accuracy in tracking	
		Dependency /Independency of working	Depending on others attention, judgement, or actions	STA2	Being checked by a colleague	
	Shared mental models	Implied understanding	Cognitively simulating the actions/cognitions of others	SSI1	Sharing information to reduce incongruence	
	Communication	Exchanging information	Identifying when information is/is not transferred	SCE1	Specifies the information that has/should be passed	
		Taking perspective of others to help with their tasks	Taking perspective of others in receiving support	SCS1	Explaining how the help could prevent further issue	
	Leadership	Setting and maintain standards	Maintaining/reinforcing/adjusting for optimal result	SL1	Domain specific practice enforcement	
		Supporting others	Directing others	SL2	Commands /asks towards safety	
		Coping with pressure	Task and team management	SL3	Goal orientated team delegation in efficient manner	

10.7 Ethical Considerations

10.7.1 Consent and Dissemination of Results

The Code of Ethics and Conduct by the British Psychological Society (British Psychological Society, 2019) were the guide in this project, and ethical considerations were addressed by the University of Huddersfield School Research and Ethics Panel (SREP; See Appendix Figure 1). The aims, methods and measures of the study were debriefed to the participants, and questions were addressed. Participants were given information about the study and were asked to complete a consent form (See Appendix Figure 10). If a participant required a change to their data use in the project, or the project itself, a review would be performed to assess if this could be implemented without affecting the experimental integrity. If it was not possible to make the changes, the participant may not be able to participate further and informed of this. Participants could withdraw from the study at any time without explanation or consequence. However, they were told that once the dissemination of the data was included into the finished project, they would not be able to retract their data from the project. Results of the study will be available to all participants only upon their request. The results may be available at Huddersfield Library for other academic researchers to view.

10.7.2 Confidentiality and Anonymity

Participants' information was confidential, and a limited number of individuals were able to access data (internal markers, and external examiner). The raw data (questionnaire scores and video recordings) were converted to a digital format, encrypted, and stored on a password protected computer, with the codes/passwords on another computer in a second secure locations. Data will be destroyed 7 years after course completion at Huddersfield University, in accordance with regulations. Data will be destroyed by shredding of all physical material, and permanent deletion of all saved electronic data. Anonymity of participants was implemented throughout the entire length of the study. No identifying or signifying data was collected; identifying (names, places) or signifying (history, hobbies) words were not asked and if present were immediately removed to create anonymity for participants. Finished work has complete anonymity for all participants. Published and/or unpublished work will make participants anonymous.

10.7.3 Health and Safety

Researcher safety was maintained by informing supervisors of any intended meetings with participants. Telephone contact was made between the researcher and the supervisors for any issues which arose. Researcher support was by email, telephone and one-to-one meetings agreed by the researcher and supervisors. Participants could communicate with the researcher at any time if they needed to, via email, meetings, or telephone. Health and safety requirements from working with display screen equipment were achieved through regular breaks (See Appendix Figure 2 for Risk Assessment form).

Chapter 11 - Results

11.1 Cohen's Kappa Results

In total there were approximately 264 IRR codings from 5 participants compared for the non-technical skill categories between the 2 raters. For the pairings of 133 codings (3 missing) there were 19 which did not agree. To note, if the percentage agreement of ((265-19)/265) was simplified and calculated, then 245/264 would produce the final percentage of 92.8%. However, the more expected chance of agreement was better accommodated by Cohen's kappa which resulted in a lower but more accurate value. Where P*O* was the proportion of codings that were rated identically, and P*E* was the expect probability of chance agreement the kappa formula was, k= (P*O*-P*E*)/ (1-P*E*). With this addition of guessed or randomly categorized data, Cohen's kappa results showed a substantial agreement between the two raters, k=0.78 (95% CI, .68 to .88), p <.001. This was translated to 78% IRR- being 15% less congruent than the percentage agreement calculation. The standard error (SE) in the measures of agreement was returned in SPSS as 0.051 which allowed the calculation of the upper and lower confidence bounds with, k-1.96×SE(k) to k+1.96×SE(k). This is using Cohen's (1960, p43) formula for standard error, $SE_k = \sqrt{p(1-p)/(n(1-p_e)^2)}$. Confidence Interval at its lower bound was 0.02 below the target of .70 as 95% of the codings fell within the .68 to .88 range of agreeability (hence the mean of k= 0.78). This was still within an adequate agreeability range to suggest that the categorizing of the coding verbalizations could be independently reproduced to a substantial agreeability level (Trickett & Trafton, 2009). However, the effects of the CI bounds and further interpretation of Kappa were discussed in chapter 11.

11.2 Use of Additional Resources and Current Use of Technology

Participants' demographic information, experiences, and technology related questions were primarily a screening tool to ensure the participants had approximate homogeneity in their training and resource usage before intervention was applied. By asking these questions, a simple understanding of how parallel their experiences were to each other was formed to assert if there were any particular measures which were not within a suitable range, in context to each question. The primary choice for 78% of participants for use of additional educational resources was online material such as YouTube videos, online videos, or mobile applications. Being the only other option chosen, participants 5 (experimental condition), 8, and 12 stated they optionally attend more surgeries as an additional resource. Furthermore, the primary use of a computer was for 64% of participants (5 in experimental condition, 4 in control condition) with tablets and smartphones being preferred by the remaining 36% (2 in experimental condition, 3 in control condition). Exactly 50% of participants stated they always use technology in their work or study- this was to an extent that they depend upon the usage of technology. The remaining participants stated they mostly used their primary technology, with 1 participant in each condition stating only some usage. Three participants from each condition stated they knew of and have used a virtual reality headset. The remaining participants knew of them but had never tried one. This may be an issue as participants that have used the headsets may have reduced effects from novelty of usage than those who have not used the headsets- this was considered further in the discussion. Overall, most participants used technologies on a routine basis and sought additional knowledge primarily by further use of technology. These results were approximately distributed between the 2 conditions.

11.3 Confidence in Instruments, Anatomy, and Procedural Steps

Being in their last year of Operating Department Practitioner training and with their experience in theatre, 85% of participants stated they agreed or strongly agreed that they felt confident about the use of instruments in theatre. However, one participant in each group stated they neither agreed nor disagree. The confidence in anatomy was evenly spread with 5 participants strongly agreeing they felt confident (3 in the experimental condition), 5 agreeing they are confident (2 in the experimental condition), and 4 participants neither agreeing nor disagreeing (3 in the experimental condition). Confidence with the operating procedure sequences in surgery were distributed equally between both conditions- as for both conditions 4 participants agreed, 2 participants strongly agreed, and 1 participant neither agreed nor disagreed. These differences in participants' confidence may be a factor in any differences in their outcomes associated with their cognitive abilities or performance on the instruments, procedures, or anatomy related elements. These descriptive statistics suggested the participants had similar clinical and educational experiences that spread between both groups. This was beneficial to the design of the experimental setup as one assumption was that trainees in the same course year should have been assessed to be on the same technical skill level. These comparable education/experience results were positive indicators towards this assumption being applicable.

11.4 Situational Awareness Global Awareness Technique Frequency

11.4.1 Overall SAGAT Comparisons

The results for all participants' SAGAT measures before and after their respective intervention group were firstly reduced to the simple quantitative changes in coding event frequency. They suggested that there was a variety of changes which had no obvious pattern upon visually examining the data. There were occurrences of being *less* accurate in the post-condition compared to the pre-condition measure (experimental condition- 27 times, control condition- 15 times), but also participants answers being *more* accurate in the post-condition measure compared to the pre-condition measure (experimental condition- 41 times, control condition- 33 times). When reduced to the frequency of changes this suggested that the participants in the experimental condition improved their results for the cues in the SAGAT towards the correct answers more that the participants in the control condition. However, this did not take into consideration the magnitude of increase/decrease accuracy.

To evaluate changes for each question more specifically, a Wilcoxon Signed Rank test was performed on the SAGAT data. As the data was ordinal it was inherently assumed to not have normal distribution due to inestimable differences between categories (Stevens, 1946). This was a factor in choosing of the Wilcoxon test along with being a matched pairs data sample (i.e. pre-post questions). In the experimental condition, the Wilcoxon signed-rank test showed that only for question 4 did the intervention material improve participants' accuracy towards the correct answer. The scores before intervention (m=2.14, sd=0.38) were less accurate compared to the SAGAT score after intervention (m=3, sd=0, correct answer= 3, z = -2.45, p=0.014). Question 4 was '*In the last ten seconds, how many trainees were in the room?*', this related to teamwork and updating one's awareness to keep track of the colleagues around. This improvement toward the correct amount did not occur for the control condition as accuracy decreased (pre: m= 3, sd= 0.82; post= m=2.86, sd= 0.69; z= -0.58, p=0.564).

Conversely, in the control condition the Wilcoxon signed-rank test showed that for question 7 the intervention material did improve participants' accuracy towards the correct answer. The scores before intervention (m=175.43, sd=16.62) were less accurate compared to the SAGAT score after intervention (m=148.57, sd=23.04, correct answer= 161, z= -1.0, p=0.027). Note that the values decrease but accuracy improved in the direction towards the desired number. Question 7 was '*As accurate as possible, how many seconds was the scenario?*'- this related to the conscious awareness of time and its passing as time perception can aid in more accurate/stronger decision-making, communications, and shared mental models (Gillespie, Gwinner, Fairweather, & Chaboyer, 2013). This improvement toward the correct figure did not occur for the experimental condition (pre: m=147.14, sd=30.94; post: m=168.57, sd= 29.69; z= -1.4, p=0.161).

By assessing the means however these results were converse to their initial proposal that the control condition was more accurate than the experimental condition. The means reflected that both groups had an average of 14 seconds of deviation in their pre-condition SAGAT, but the experimental condition increased their time by 21 seconds, resulting in being the closets to the correct time. Whereas, the control condition underestimated time in the post-condition measure. This suggested that in 2.5 minutes of material the experimental condition participants had 4.3% deviation of time overestimation whereas the control condition group had 8.1% deviation of time underestimation.

Between groups differences may have validity due to the non-significant difference suggested from a independent-samples t-test in time estimations before intervention (t (12) = -2.134, p= 0.054). This was discussed in section 11.7.

As question 9 was at the nominal data level with pre- and post- data comparison, a descriptive exploration of the conditions showed that the intervention material improved participants' selection towards the correct answer. Question 9 was '*what colour were the three bins?'* and was a distractor question as it had no meaningful relevant to patient safety and were not used as cues in the scenarios. Before intervention most participants stated they did not see the items (experimental= 72%, control= 86%). The incorrect colour yellow was given (experimental= 14%, control=14%), and the correct colour orange was given from 1 participant only (experimental= 14%). This suggested their attention was not being deployed to these items as their information gathering processes have no purpose for them in regards to the various non-technical skills occurring in the scenario. After intervention, some participants stated they did not see the items (experimental= 14%, control= 72%). The incorrect colour yellow was given (experimental= 28%, control=0%), and the correct colour orange was given from 4 participants in the experimental condition (57%) and 2 participants in the control condition (28%). This suggested that a consistency was held in most of the control condition participants who continued to not deploy attention to the items. However, there was a change in accuracy for the experimental condition which suggested the majority did deploy attention to these items. The cause of this difference was considered in the discussion section.

The above exploration of participant's data required support from in-depth qualitative analysis. If simple increases in higher level skills such as *prediction* and *transformation/reflection* were the goal, then qualitative analysis would not be required. However, a detailed exploration of *how* cognitive processes were being modified in each condition was essential, to explore more than if they increased or decreased. The frequencies of NTS related cognitive processes as represented by cognitive verbalizations, have more meaning when provided with additional context from the qualitative analysis for each participant. Certainly, within a single session of either theory-based material or practically driven scenarios, each participant's cognizance changed. However, without the context of the participants' cognitive verbalizations the SAGAT results had less power to explore the effects of both intervention materials. As cognitive verbalizations were later analysed, the SAGAT results therefore provided more context for individual improvements from their respective condition.

11.4.2 Distractor and Researcher Bias Questions

Regarding the distractor questions, question 9 was referring to an object that was in the environment but not relevant to any non-technical element of interest. Post-condition measure answers for 4 out of 14 participants changed from the incorrect to the correct answer. This suggested that these 4 participants chose to deploy their attention to the cue in the post-condition measure to be able to answer the question correctly. Question 24 was a distractor in that it was an action of the surgeon that did not have any meaningful input to the participants goals during the scenario. However, some participants did identify this action in the pre-condition measure, meaning that they identified and remembered this action before any mention from the experimental material was presented to them. Thus, it was possible that there was a meaning to this action that the researcher did not account for and was elaborated on in the discussion. However, the 2 questions 'planted' to assert if participants showed bias did not receive any biased answers. The 2 questions were created with reference to items in the environment that *did not exist*. If bias were shown it would have been easy for a participant to simply guess the answer in hopes of being correct. All participants stated they did not see these items in both pre- and post-conditions suggesting that they were honest in their replies which implied they did not show researcher bias.

11.5 Decision-Making Strategies

11.5.1 Option A

The decision-making scenario results (see Table 2) showed that 10 out of 14 participants (71%, 6 in the experimental condition) ranked option A as their 3rd choice, therefore the least desired choice of action towards the scenario provided. Three participants in the control condition selected option A to be their 2nd rank choice, and 1 participant in the experimental group selected option A as their 1st choice. This meant that 71% of participants identified option A as the least desirable choice for the optimization for patient safety. With three choices the distribution for ranking should be 33% per option. However, this result suggested some form of purposeful discrimination against option A as more participants selected this as the worst rank even though the 3 options were contingently equal. Option A was ranked 2nd by 2 CAPs and 1st by 1 CAP- therefore in agreement with only 4 participants. There was a counter-intuitive description whereby the details suggested that the surgeon should switch tools, but due to the surgeon stating he was not switching tool then a rule-driven and hierarchical theme was provided. Most participants agreed this strategy was not the best for the scenario described, which indicated they may break the hierarchical and rule-based dynamics if it was towards patient safety.

11.5.2 Option B

Option B was more analytical in the approach as the option revolved around the increase of people in the OR to gather information and understanding, along with increased behavioural abilities to make a 'better' decision when compared to the participant's individual analysis of the situation. This appeared to be a team-based approach, and the process to further analyse the environment was embedded within the option. The main difference to note was that the 3 CAPs who also previously ranked and explained their decisions all marked option B to be their last choice. This meant that participants 2 and 8 were the only trainees who were concurrent with the CAPs. For option B, 10 out of 14 participants (71%) ranked this as their 1st choice action for the scenario presented. Two participants (1 in the experimental condition) ranking it their 2nd choice. The explanations from each participant may highlight why this contrast in choice occurred, and reasoning behind the difference in choice for participant 2 and 8 was later explored.

11.5.3 Option C

This option was orientated towards a recognition-primed decision-making strategy as the answer was based on the identification of key cues that indicated the surgeon was experiencing issues. These surrounded immediate intervention being required before the status became critical. Option C was identified as the primary choice for 2 out of 3 CAPs and 2nd choice for the 3rd CAP. For option C, participants 5, 7, and 8 selected this option as first choice and agreeing with 2 CAPs. Ten participants (7 in the experimental condition) ranked option C as second choice, with 2 participants in the control condition ranking the choice last. This summarized to suggest 85% of participants indicated option C was not the best rank to select towards patient safety, which again opposed the CAPs assessments.

Table 2

Table showing the answers to the written decision-making scenario. Option B was contrariwise between most CAPs and trainees.

Participant	Option A Rank	Option B Rank	Option C Rank
CAP 1	2	3	1
CAP 2	2	3	1
CAP 3	1	3	2
P1	3	1	2
P2	1	3	2
P3	3	1	2
P4	3	1	2
P5	3	2	1
P6	3	1	2
P7	3	1	2
P8	2	3	1
P9	2	1	3
P10	3	1	2
P11	3	1	2
P12	3	1	2
P13	2	1	3
P14	3	2	1

11.5.4 Decision-Making Inferential Results

As there was variation in participants ranking orders, a Friedman's Test (Friedman, 1937) was performed on the data to find any statistically significant differences in these decision-making choices depending on the option response (See Appendix Figure 17A). A Friedman test was selected as the assumptions of normality were not met by this small 42-point (14*3) dataset and is a non-parametric alternative to the repeated measures one-way ANOVA. Additionally, there were 3 ranking measures all at the ordinal data level which met the assumptions of the test. The test compared the mean ranks between the related groups to indicate if they differed. The results showed that in the 14 participants there was a statically significant difference in the type of strategy applied depending on the NTS options which were deemed most suitable to solve the scenario provided (X^2 (2) = 10.43, p= 0.005). The Friedman's Test is an omnibus test and did not identify which option was ranked differently from another. Therefore, a post-hoc analysis with Wilcoxon signed-rank test was performed on the data with a Bonferroni correction applied. The results suggested that Options B and C had more consistently ranked order between them therefore did not have significant differences in ranking order (Z = -1.41, p= 0.152) therefore Option B was ranked 1st choice (m=1) most consistently and Option C 2nd choice (m=2). However, Option A did have a significant difference between being selected as the last choice over Option B (Z= -2.57, p=0.01), and Option C (Z= -2.24, p= 0.025).

11.5.5 Quantitative Data Summary for Decision-Making

Overall, there were 6 possible combinations of ranking order- 5 of these outcomes occurred within participants. This indicated that within a sample of 14 individuals with similar practical skills as approximated by the academic year, decision-making abilities were significantly different in places. The decision-making ranking orders from participants suggested that when measured by a written scenario with clearly defined sets of information into 3 options, 57% of participants followed the ranking order of 3,1,2. This translated to suggest rank 1 was option B, rank 2 was option C, and option A last -option B was the more analytical and team-based approach. This layout of results suggested the trainee participants have similar strategy based on the frequency of the 3,1,2 ranking order. From this, a more individualistic prediction of the type of cognitive processes used in each intervention condition was required to allow more understand of each participant's capability prior to intervention. The uncertainty of option C being the last rank indicated that participants were not cohesive in agreement that waiting for the surgeon without interference would be the best approach for patient safety. Yet, option C had interruption as part of its answer, but also was not the primary choice. Therefore, a factor(s) within option B were more valuable to the trainees than the other 2 answers, yet was the least valued by the CAPs. When presented with the same scenario they identified, interpreted and assessed, then selected their final rank order with a heterogenous outcome. Certainly, most ranking orders were not the same as the CAPs - the captured cognitive verbalizations were explored in section 10.5.

11.6 Individual Decision-Making Strategy Exploration

11.6.1 Significantly Occurring Ranking Order (3,1,2)

For the 8 participants (57%) with this ranking order, individualistic exploration was performed on their cognitive verbalizations to recognise how and why they used NTS based cognitive strategies to conclude upon the resultant order of choices. Experimental condition:

Participant 1- Option B (1st) was regarding improving the situation oriented towards a team-based and information gathering approach. This opposed the CAPs answers who selected this as a tertiary option. The participant's underpinning non-technical occurrences oriented around the goal to source more information, perspectives, and generally more assessment of the situation before a decision. Participant 1 stated, '*It's just always better to have more people if you are on your own, because you are on your own, even though there is a surgeon and an anaesthetist you are on your own. 'Cos they don't do stuff, they expect you to do it* (I2)'. With the tacit goal to obtain more information, the participant's elaborations suggested they desired a second opinion before deciding upon an action. Being alone itself to select an action would not be an issue, as a decision could be made easily if regardless of its consequence. However, it may be the inability to seek a second opinion before decision that lead the participant to select this option- as the other options A(3rd) and C(2nd) would depend on only the participant's judgement. This primary choice selected imparted that further information gathering was needed before selecting an option to perform, which may be why it was not an optimal choice for the CAPs as their opinion would already be most accurate in comparison.

The participant's 3rd ranking of option A was last choice as they stated this option would cause them to break a rule (being to leave the OT). They elaborated to state '*Because you can't just leave the room, if there is something like that going on in the room you can't just leave the room'… 'I don't know about that one, as you wouldn't leave the room but then you need that equipment.* (I4*)'*. When analysed along with their first choice, it could be suggested that the participant's decision strategy was influenced by rule-based practice. They did not select option A(3rd) due to breaking of a rule, even though they identified this was necessary. Additionally, the selection of option B(1st) may have suggested they may need either verification or some form of aid from others to make sure they have, and were allowed to, make their choice (for example if option A stated the surgeon gave permission to leave the room). Therefore, as decision-making has been theorized as a resultant process of various factors including the participant's situational awareness levels, it provided support that their perceptual and comprehensive skills may also be at a rule-based level prior to the intervention in this thesis.

Participant 3- There was not a sufficient number of cognitive verbalizations for participant 3 to explore their cognitive strategies in relation to the differing decision-making models. To avoid biases, the data gathering method did not include prompts during participants verbalizations therefore if a participant did not talk, the researcher could not probe them. Participant 3 provided 'to the point' answers without elaboration. This meant that with the absence of decision-making strategies used the predicted types of NTS used could not be deliberated and later insight of the changes in these skills may not be as exploratory. All 3 CAPs selected option B as the last option with rationale that with more people there '*may be more chance of confusion*' (CAP 3, I9). Participant 3 acknowledged the presence of staff but chose to include more, she stated '*There is already other people there, other staff. And then I'd go get more staff, like another ODP or something I think I'd go for that one first (I1).* Based on this, option B(1st) was selected as it allowed improvement of understanding by a team-based and information gathering approach. The rationale has contrast to the CAPs however nothing more was able to be understood as to why participant 3 believed this would best improve the situation.

Participant 4- Participant 4 produced the ranking order of 3, 1, 2 in alignment with 57% of participants, but with a novel explanation for option A(3rd). Participant 4 stated that there was a benefit to allowing the surgeon to continue with the current actions, and that giving a little more time may help the surgeon to verbalize the issue better. This was an analytical approach to the situation as the scenario detailed that a significant amount of time had elapsed and was a counter-intuitive strategy. However, the rationale for ranking this option the last was because to *'have a backup ready is ok but you're not really doing anything*' (I4), therefore option A(3rd) was too passive for the scenario. This meant it may not be common issues such as the hierarchical and rule-based consequences, but that *no action* from the trainee was the least desirable aspect of this option.

Participant 6- The primary choice B was last ranked for the CAPs, and oriented around a team-based and further information gathering approach. There was indecision between option B(1st) and C(2nd). The participant identified that the surgeon's perception of their own skill may be overestimated, e.g. '*he might just be out of is depth and thinking I've got this like I'm a surgeon thinking he's a bit big-headed'*(l5). However, the participant chose Option B(1st) as the additional help from others allowed a safety net or a failsafe procedure if intervention was required. Participant 6 described her thoughts that *'…you need more staff if CPR is needed or if you need to start lifesaving treatment*' (l5). This may have logic contradiction in that if the surgeon was interrupted, further critical support may not be required as the issue may have been addressed. This rationale of the participant may suggest they may not want to interrupt the surgeon, regardless of their own judgement. Explicit verbalization of this was also evident as participant 6 stated, '*…it seems quite demanding like' you should do this you should do that'. At the end of the day you are an ODP they are professionals in there. It's just it's a bit much'*(l14).

Although this team-based strategy was selected, the rationale for option C(2nd) was more analytical and considered than option B(1st). Like CAP 1 and CAP 2 who selected option 3 as the first rank, the participant identified 2 critical cues they used to justify the reason for calling a resuscitation team. One cue was that the throat was time-dependant and the other cue was that there was enough staff in the room already to be able to leave the room temporarily. They stated, '*You will have staff there and you could quickly run and I thinking getting the resuscitation team is a good idea because it's been 7 minutes now where his throat's been blocked'*(l5). This demonstrated analytical thinking based upon the cues they identified as important in the scenario, with the CAPs mentioning these 2 cues also. In addition, they explained the possible consequences of these cues *('stats would be down, he wouldn't be getting any oxygen to the brain'*, l10) to aid them in predicting what would occur next. This demonstrated level 3 situational awareness. Yet the mentioned 'safety net' and hierarchical perceptions within the OR demonstrated their influence to the final choice of option B as rank 1.

This rule-based perception of hierarchical status was expressed under the conditions of the scenario, as the participant undermined their own judgement due to their role of an ODP and emphasizes the judgement of the surgeon, comparatively labelling them 'professional'. This suggested the participant may view a limit to the value of their judgement - even though the scenario given suggested an issue with the surgeon's own judgement. It was also possible they believed they were missing some information which the surgeon may possess to form a different decision. Overall, the team-based choice of option B(1st) for participant 6 may be due to indecision and hierarchical perceptions which were the influences to selecting more people to help. Overall, the participant has critical analytical strategy but was mostly socially orientated, therefore in subsequent measures more socially orientated verbalizations were predicted to be produced.

Participant 7- Participant 7 also did not provide any evidence for rule-based decision making, but elements of recognition-primed and analytical strategies. This helped to come to an apparent logical elaboration for the rationale to rank option A(3rd) as last, instead of 1st or 2nd choice (CAP3 1st choice; CAP1 and CAP2 2nd choice). Participant 7 recognized that the surgeon may be experiencing attentional bias by being focused on a specific task, and further summarized the option to state '…*the surgeon keeps saying he 'nearly has it…'(l3).* Then, he recognized that the patient status was within severe risk -the question indicated that '*the status is about to become critical'*. However, the patient made an analytic assumption based on these two pieces of information. Participant 7 did not assume that the new state of the patient being 'critical' would alter the perceptions of the surgeon. This assumption- *that the surgeon would fail to identify the critical status of the* *patient*- had no obvious basis apart from the information provided in the scenario details that the surgeon was experiencing an issue. Adding such indirect/implicit assumptions into the decision-making process required the ability to comprehend the general cognitive state of the surgeon, and to have forethought into how real-time events can change based on a decision. This logic was similar to CAP3, as CAP3 also supported that '…*it says right there 'patient status is about the become critical'. Which tells me it's not critical... let's assume they do, go critical.*'(l3-4).

However, the deviation in final selection between participant 7 and CAP3 may be that CAP3 inferred the surgeon would acknowledge the situation and agree to the tool usage, whereas participant 7 inferred they would not acknowledge the situation. CAP3 justified this to '*give the surgeon a chance by the time I am back he either has or has not, either way I am ready for the next part and we are all on the same page of using the tool.'(l7).* However, participant 7 stated '…*but it can end up being 15 minutes and the patient's critical and they haven't realized…'(l4)*- the implication that 15 minutes has gone by would suggest the issue still persisted with the surgeon, and that a shared understanding of the situation had not occurred. Although different to the CAPs, participant 7 demonstrated strategy that was had rationalized analytical observations. Based on this, it was predicted there should be more frequent occurrences of more complex cognitive verbalization and less basic perceptual, or simpler rule-based strategies for participant 7.

Control Condition

Participant 10- There were 7 teamwork and communication coding occurrences which described the rationale of the participant towards selection of option B(1^{st}). The rationale was that as the surgeon was having issue, the solution would be to find assistance from individuals other than the participant. Indeed, an assumption about the meaning of more people was explained as 'i*f you feel that the surgeon requires more assistance then it's better to get more staff. More staff equals more knowledge'*(l1). This assumption was rapid and a shortcut-like decision, yet with more consideration they may have had similar elaborations to CAP1. For example, CAP1 stated '*I have no idea what time or day this is, but if you're going around looking you could end up with 5 more trainees, or yourself who can be a visitor in scrubs* [referring to the researcher]*, its good if it's an ENT or an otolaryngologist but its pot luck* (l8-10).

There may be the need for a second opinion as the participant may want more perspectives before deciding, but such in-depth elaboration similar to CAP1 was absent. For example, she stated *'But at least if you get more people in the room, the chances are that there will be someone with closer experience and might be able to better judge the situation'(l7).* This highlighted the implicit nature of factors that can influence decision-making strategies used, as there was a perceived gap in knowledge between themselves and the surgeon which caused the participant to favour option B(1^{st}). The absence of corresponding previous experiences of a similar situation may be a reason for this cognitive difference. With a significant amount of experience CAP1 opposed this reasoning to find more people as it may take too much time. He stated '*Sometimes you ask for a tool and people are gone 15 minutes and they have been trying to find one so it's a risk'*. Therefore, due to previous similar events CAP1's comparative analytical and predictive strategies were more developed and a reason why this was the last rank for all CAPs. It was expected for participant 10 that there would be specifically team-based cognitive verbalizations such as how each person in the video materials were supporting each other. Also, the participant may verbalize how and when information was shared to form a shared mental model. This was based on the participant's verbalizations explaining that more help can be '*a 3rd perspective who just comes into the room and has a fresh look'(*l8).

Participant 11- The rationale for selecting Option B(1st) was less socially orientated but surrounded the identification of tool usage. She stated, *'like the tool is nothing new and it says its standard practice. So, for whatever reason the surgeon is not using that tool' (l2).* The participant then attempted to take the perspective of the surgeon to understand a reason why they may be continuing with their initial attempt. The participant stated, '*so he must be aware of the tool so he must have already thought about it maybe. So even if I just make eye contact with the surgeon and say about it I bet they won't use it'(*l5). Although a short explanation, this had complex steps in non-technical understanding when broken down. It was encoded as level 3 situational awareness as the participant had to understand the basic elements of the scenario (perception), understand the links between each component that made up the scenario- (comprehension), and how the resultant choice will project in the future ('*I bet they won't use it'-l6).*

The positive viewpoint of option B(1st) was then stated which was similar to other participants when anticipating the help of other more experienced individuals. The participant explained they were '*thinking option B to be honest as someone, hopefully another surgeon could do it and interrupt maybe (l6).'* This strategy had rule-based elements but were more orientated around the participants viewpoint that the surgeon had already considered other options, therefore the reason why this action continued. The only way to change the surgeon's behaviour was with comparable surgical staff members. Lastly, although option B(1st) was the final 1st choice, the participant did show hesitance. She explained '*I'm not sure what the right answer is here, if i was there and it felt right then yes, I would at least make sure they heard me even if it's just a suggestion but for this I'll pick B*…(*l7).*' This may have occurred because there were positive elements to all options, and to force a decision may have inaccurately captured this participant's desired actions. There was evidence of analytical mindset to evaluate each option and take perspective of the surgeon to further understand the cues in the scenario. Yet rule-based and role orientated constraints may have been influential on the participant's final choice, and similar cognitive verbalizations were therefore predicted in the pre-condition measure.

Participant 12- With a reoccurring theme parallel to many other participants, the dependency on more individuals to be secondary perspectives was described by participant 12. She explained that more people brought into the situation meant that, '…*it's a group effort to think about what's happening, you have more people to compare it to, to think right is this normal, is this what we would be doing if we were doing it or should he not keep trying*(l2)'. This had the underpinning consideration of the group's shared awareness of the situation to help each other understand the events unfolding. The participant therefore demonstrated investigative assessment instead of suggesting immediate rule-based procedures.

The cognitive processes of taking the perspective of others was included again in the participant's decision towards option B(1st)- she stated *'you're not leaving the surgeon to feel they've been inadequate in the situation you're still giving them a sense of that you have confidence in them* (l4)'. Again, each course of action was deliberated by the participant for its potential effects, and the conclusion was based around how much trust they have in the confidence of the surgeon in their decision to continue. E.g. she stated *'I think that's what it's all about really the confidence of the lead surgeon in maintaining the safety of the patient* (l7).

Finally, at the end of the participant's transcript she showed signs of creativity and explained why she would merge two options. This demonstrated the ability to utilize information into new combinations that may not appear obvious but would be a better solution. She explained '*The resus team in option C is good I would pick that part out of that answer, so kind of half B and half C. But B would be a really good way to not erm, jeopardise the rest of the theatre'.* Therefore, participant 12 demonstrated creativity by weighing up the consequence of each action, and determined that the surgeon's confidence could be impaired if he was interrupted- this may have greater effect on patient safety and a knock-on effect to the team. It was predicted that this participant's pre-condition measure was patient-safety orientated as the underpinning influences were driven by error prevention. This meant they may have already had decision making strategies which were at a level desired in the goal of this thesis.

11.6.2 Summary of Findings

There were commonalities in reasoning for the selection for option B within these 8 individuals, and were socially-orientated over other types of non-technical abilities. More staff were understood to bring an array of more opportunities and information that could change the unfolding events. However, with deeper analysis of their cognitive verbalizations a large variation in the information identified was noted, along with different weights and interpretations of importance attached to them. These also mostly differed from the CAPs. The result was that participants deviated from each other's rationale due to less/more developed decision-making methods but all had a universally coherent mindset of ranking order 3,1,2 which they believed was best for patient safety. Participant 12 especially showed higher ability to decipher how best to approach the scenario even though their decision-based cognitive verbalization encodings were not as frequent. This exploration allowed better context of the later measures of NTS usage, as changes in cognitive verbalization frequencies and NTS perspectives (e.g. from rule-based and team dependency to increased foresight with independent action) can be contrasted greater due to the intervention material presented.

11.6.3 Non-Significantly Occurring Ranking Order (3,2,1)

Occurring within 2 (14%) participants, this ranking order did not match most participants as 1st and 2nd place ranks were swapped. However, these 2 participants matched their first choice to 2 CAPs which suggested their strategies may be further developed, or they had better ability to verbalize their cognitions during session. Exploration for each participant was performed to find why this NTS category differed from the same cohort.

Participant 5- When asked to verbalize her choices participant 5 displayed signs of being unsure and after re-reading the scenario still did not verbalize rationale. The researcher did not prompt the participant but improvised by asking a second 'backup' scenario (not included for other participants to reduce session time scale, see Appendix). The logic behind this was that the rankings provided could be assessed to see if the same strategy was primarily used for both scenarios. Even though they had different contexts the participant could have picked out the options which both have the same embedded decision and NTS related processes. This could also be compared with the CAPs to find any matching order but did not have the walkthrough of their deliberations. The participant selected a less common ranking order of 3, 2, 1 for the first scenario. This represented more immediate procedural action followed by team based and analytical strategy. However, the participants ranking order was 2, 3, 1 for the second scenario which suggested a different type of decision-making strategy for this scenario. This suggested that context may affect the type of strategy used and had limited meaning without the inclusion of cognitive verbalizations to help investigate rationale in more detail. This highlighted the importance of cognitive verbalizations and a potential barrier was preventing participant 5 from providing these.

Participant 14- Participant 14's decision making for options A, B, and C were ranked in order 3, 2, 1. CAP1 and CAP2 also selected option C as 1st rank, with CAP3 ranking it 2nd. The cognitive verbalizations were similar to participant 8 who also selected option C as the first choice (see below). It was similar as there were several deliberations and expressions of being unsure about the choice. Participant 14 stated, '*option A I've put down as the worst even though its similar* (to option C) *because the situation is quite bad but that sounds passive* (I5)'. There were also phrases such as '*I'm not sure (I4), 'I don't know really(I6)', and 'I'm not sure if I would (I8)'* even though the decision-making outcome was closer to the CAPs and less like the commonly occurring 3,1,2 pattern.

There was consideration for each option that reflected similar thoughts from the CAPs. For example, she stated, '*B is good, but I was thinking what if its early in the morning and there are less people around I don't think a cleaner can do much* (I7). This segment of information (to get more staff) was rapidly rationalised by most participant to be beneficial as more staff inferred more skill. However, participant 14 was the only individual to exactly match the logic and larger contemplation of the CAPs. CAP 1 stated '*I have no idea what time or day this is, but if you're going around looking you could end up with 5 more trainees* (I8)'. This was not expected in the cohort of participants in this study and highlighted the wide variety of NTS capabilities in the same group. Some NTS skills show calibration in participants, yet others do not. This meant the full range of decision-making levels have been demonstrated in this sample- from simple rule-based practice through to expert-like analysis of information.

11.6.4 Non-Significantly Occurring Ranking Order (2,1,3)

Occurring within 2 (14%) participants, this ranking order had option B as the 1st choice. Again, this was the 3rd choice for CAPs, and it was predicted the rationale for participant 9 and 13 would be the same as the others. However, options A(2nd) and C(3rd) were swapped therefore there was interest to find how and why these 2 participants chose the CAPs 1st rank option C, as their last choice.

Participant 9 -The decision making for options A, B, and C were ranked in order 2,1,3. For the last ranked option C, there was an element of personal culpability and job role limitation which influenced the decision-making strategy used. A rule-based perspective was immediately stated by participant 9: *'That's not your job. You shouldn't be telling them to use that tool. You don't say that, it's up to the surgeon what he uses.'(l2).* Prediction was then provided for this boundary of task delegation: '*if the patient does die and they used the tool, the reason why he died was because you said to the surgeon 'by the way use this tool,' you're held accountable for that'(l2-4).* There was then an awareness of this rule/role orientated limitation, as the participant then deliberated further. She stated '*I think it comes down to hierarchy of whether you want to speak or not. So, f you're entirely sure they should use that tool, then tell them to use it by all means.'* This suggested there were circumstances which allow this initial rule to be broken. This accountability-based consideration was only mentioned by participant 9.

A change of strategy occurred towards the end of the participant's cognitive verbalizations as she explained why option B was the 1st ranked choice. She was more analytical in the expansion to state- '*But for option B you can get more people in and if the surgeon is still trying then we can all say to them, 'now it's time to try the other way'(l8).* Noteworthy was the use of '*we*' which suggested the participant perceived a group effort in asserting to the surgeon to change behaviour. This had similarity to the 'safety net' reasoning displayed by participant 6 in that team support may be able to intervene with the problems more effectively due to a collective effort.

Although the scenario implied the surgeon was facing attentional bias, the participant held this information as an indicator the surgeon should be left to concentrate. She summarized that '*I'd still says it's their choice it would be hard to say to them if it's at a critical time, as I don't want to put them off, and end up being wrong. If someone else thought the same then I'd maybe suggest it, so more people would help decide'(l10).* Therefore, several socially based factors were components in the participant's ranking choice, which was an indirect goal towards patient safety when compared to the other 2 more immediate options. These non-technical elements were predicted to be apparent in the participant's pre-intervention data collection, in that some analytical appraisal may occur but rule and socially orientated verbalizations would be the prominent occurrences.

Participant 13

The same rationale for option B was present as the dependency on others in providing more information was verbalized, the participants stated, *'If just the one person is off their game, isn't concentrating, then it could affect everyone else. So, I would get more people as they can see the situation and help vote what's the best thing to do* (I6)'. This strategy is not as time dependent and may use it there is not an appropriate amount of prediction of outcome for a choice.

This team-based approach was also included in the rationale for choice of option C(3rd) as more individuals were assumed to be able to change the surgeon's problematic behaviour. I.e. *'And to be able to support each other and to think right how we can support this person so we can get through the surgery the best we can. So, if that means the surgeon needs to be told he needs to switch the approach then you should* (I13).

Participant 13 did not include predictive cognitions or more in-depth deliberation past this perspective, however the key decision on option C(3rd) being ranked as the worst choice was application of a real-life context to the written situation, she stated, *'but in theory and in practice its different as I don't want to be wrong if I interfere, even when the patient would go critical then that's beyond my control, so I'd not choose this one* (I16)'. The underpinning rule-based consequences of option C in real-life was a primary reason this was ranked last even though participant 13, and some other participants, had more elaborate and patient-oriented interpretations for this choice.

11.6.5 CAP matched Non-Significantly Occurring Ranking Orders (1,3,2 and 2,3,1)

Occurring within 2 (14%) participants, the 3 options were ranked in the exact order of the CAPs. There was assumption this indicated the highest levels of NTS were used such as attempt to predict events, creativity, and transformation of information. However, with previous variation of results this may not have occurred therefore individual exploration was performed.

Participant 2 - The choice of decision making for participant 2 followed a different ranked order of 1,3,2 from 76% of participants and a different order to CAP1 and CAP2. However, the 3rd ranked option B(3rd) of participant 2 matched the 3 CAPs answers. The ranking order matched CAP3 which suggested participant 2 had less difference in decisional strategy to the CAPs compared to 76% of participants. When analysing the explanations for the choice of participant 2, they immediately express how an increased quantity of individuals did not reflect an increased quantity of help. She stated '*It is very passive as its just adding more people to the situation and hoping someone knows how to solve the problem. But we already know the problem'* (I2). This rationale was more consistent with an analytical strategy of decision-making as it had elaboration of understanding past the initial concepts presented (i.e. more people will help the situation improve). This systematic logic was suggested when the participant assumed the perspective and skillsets of others in their information gathering process. She stated '*we know the surgeon is having issues with the current method so getting another instrument that can work is a lot better than more people. I imagine the surgeon knows how to use the tool he has but something isn't right still'* (I5-7).

The analytical strategy was also highlighted in the rationale for selecting option A(1st) over option C(2nd)- in option A there was no urgency but in option C there was urgency to use a certain tool. The participant's rationale for this was a very similar perspective as CAP3. CAP3 explained '*It all sounds bad on the surface, but it says right there 'patient status is about the become critical'. Which tells me it's not critical. Which tells me I don't need more people in the room'.*

This demonstrated how each piece of information provided in the scenario can be interpreted with more 'weight' or importance for each individual. This same piece of information held a high importance for the participant as she stated '*the scenario says it is 'about to become critical' so if I'm ready with the other tool, then the next downhill sign I spot would probably move this option* [option C] *to rank 1*'. This suggested that although this participant was in the same technical skill range as the others by measure of year group, their decision-making ability had similarity to more experienced individuals. This was considered when analysing the remainder of participant 2's results as their cognitions may have complexity that could be overlooked upon interpretation.

Participant 8- This choice of decision making for participant 8 followed the same order as two of the three CAPs, being 2,3,1. This order was not evident for any other participant therefor atypical to 93% of participants. As the ranking order matched 2 CAPs, at face value this suggested participant 8 deployed the most similar decisional strategies to the CAPs compared to all participants. However, assessment of their verbalizations to understand the underpinning occurrences was required to explore if these similarities were as developed as they were anticipated to be, for the ranking order chosen.

There were occurrences of considerable higher-level situational awareness by the participant- indicated by their deliberations and justification into 'what could occur' within the provided scenario. This suggested the participant could conceptualize the written scenarios and the dynamic links between the factors of interest in the scenario. For example, participant 8 verbalized a cognitive walkthrough of events from one example: '*it's like when you're just waiting for the bleeding to stop, you don't know what damage has been caused…and they've got an airway in, you don't know what could have happened, your airway could have been dislodged'*(I26). Although there were 7 coded occurrences of similar comprehension based situational awareness cognitions, most non-technical related verbalizations fell into the decisional categories. This should not be surprising given the scenario was decisional, but the type of strategies deployed were markedly different when compared to the CAPs upon exploring their cognitive explications. The CAPs displayed unambiguous and dualistic details to their choices, whereas participant 8 had more indecision (similar to participant 14 as previously mentioned).

For example, CAP 3 immediately stated that '*Option A is better than C because why would I want to change something that is not yet so bad*'(I1), yet participant 8 immediately expressed '*That was really hard. That fine line between all three like literally fine. Fine.*' (I1). This suggested that although they achieved the same answer as the CAPs, participant 8 had more effortful deliberation evident from these frequent articulations and continuous development in understanding. For example, she described that a certain choice was *'…originally going to be my number 1, but then to continue with the procedure, and now I'm like err, maybe let your anaesthetist take over'(I24).* These glimpses of cognitive walkthroughs showed how her technique to assess the scenario became more accurate over time.

One explanation, which also could be applied to the CAPs, may be that experience of similar situations allowed them to mentally form a catalogue of previous situations and outcomes that can be revisited to compare with the current issue. Then, the previous 'closest fit' solution that worked in a previous environment could be adopted to suit the current environment. The CAPs likely have stronger abilities for such walkthroughs, however participant 8 was able to come to the same conclusion given more time. Based on the cues relating to safety of the patient she stated *'…communicating with your anaesthetist was kind of the top priority for me at that point, again it was a really fine line between them three, that was really hard' (l33).* Participant 8 verbalized enhanced recognition-primed, rule-based, and analytical strategies which suggested they had reduced variation in decision-making abilities towards CAPs than most other participants, and a different style of walkthrough to form their conclusion. The effectiveness of either material in the control or experimental condition may not be optimised for participant 8 and the considerations of their potential negative effects was considered.

11.6.6 Decision-Making Strategies Summary

Both qualitative exploration and quantitative assessment were used to analyse each participant's decision-making utilisation when presented with a written scenario with 3 differing answers. These answers had underpinning NTS themes and decision-making strategies which were reviewed by the 3 CAPs in their differences in approach. Overall, a significantly occurring pattern of strategy was used for 57% of participants which positioned towards team-based information gathering and behaviours. With division of ranking orders for the remaining 43% of participants along with minor variations in how each participant interpreted the scenarios, there was widespread variation in skill.

With the interest of patient safety and prevention of undesired events, these suggested variations were useful when later exploring any effects of intervention towards the standardization of these NTS skills. NTS abilities may have less variance around a common mean level of skill for one cohort, such as the participants in this thesis. The control condition material could better reduce variance by explaining the underpinning principles for undesired events, individual differences in performance, and the importance of NTS in general. Perhaps practical material could better reduce variance by providing consistent feedback that steers a cohort towards the same understanding of cues in real-world contexts. To identify if a similar pattern occurred for the NTS category of situational awareness, the next section similarly explored how and why participants were able to perceive, comprehend, and predict cues from the equivalent pre- and post- material used in each group.

11.7 Situational Awareness Coding Comparisons

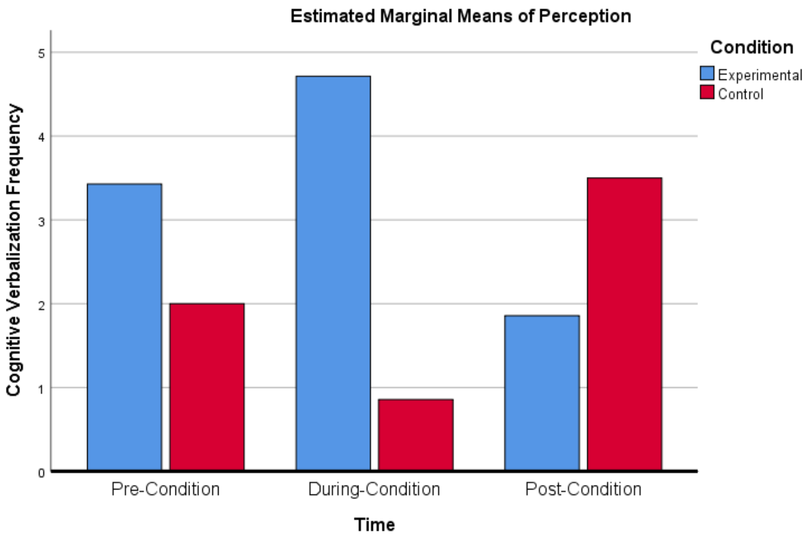
The results of all participants' coding schemes were added to SPSS and descriptively explored. There were more than 20 bottom level codes - approximately 5 codes comprised of an NTS category of either leadership, situational awareness, decision-making, or teamwork and communication. Situational awareness codes were combined into their respective category and combined when necessary. To describe this differently- the codes were combined to their NTS category, but also combined if they were derived from the same classification. For example, in the coding table CSPE1 and CSPE2 were both derived from the situational awareness category of *perception* therefore combined. Therefore, presented below is the descriptive exploration of key NTS concepts defined by previous research, which were extracted from participants' cognitive verbalizations. All transcripts before interventions can be found in Appendix Figure 12A, transcripts during intervention are found in Appendix Figure 12B, and transcripts after intervention are found in Appendix Figure 12C.

11.7.1 Situational Awareness - Comparison of *Perception*

The differences in cognitive verbalizations for cue *perception* from the pre-condition to post-condition measures showed a decrease in the experimental condition (3.3 to 1.9, -1.4) but an increase in the control condition (2 to 3.6, +1.6). This meant the conditions conversely changed the number of verbalizations for cues with basic perceptual information (e.g. '*there is a swab on the table'*). A decrease in *perception* verbalizations suggested decreased basic situational awareness that may have improved to hold information more developed than basic observations. An increase may suggest the participants continued to have limitation in developing their cognitive verbalizations therefore reproduce a similar level of basic descriptions of relevant stimuli in the scenario. Figure 15 highlighted the amount of perceptual content produced during presentation of the material in each condition- there were more than 4x more verbalizations on average in the experimental condition than the control condition.

Figure 18

Graph showing the mean frequencies of coded verbalizations over the course of the data collection period for Perception.

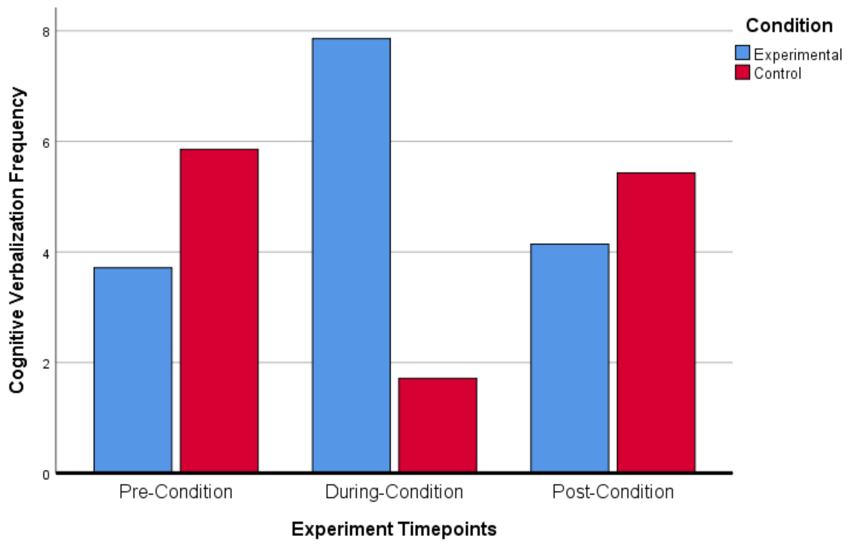


11.7.2 Situational Awareness - Comparison of *Comprehension*

The differences in cognitive verbalizations for cue *Comprehension* from the pre-condition to post-condition measures showed a decrease in the experimental condition (3.7 to 4.14, +0.44) but in increase in the control condition (5.9 to 5.4, - 0.5). This meant the conditions marginally and conversely changed the number of verbalizations that demonstrated a more in-depth understanding of the processes in the simulation (e.g., '*She should write that information on the whiteboard, and now she is doing so*'). Although there were more than 3x on average the amount of comprehension related codings occurring during intervention in the experimental condition (see Figure 19), this did not translate to having a notable effect when compared to the control condition.

Figure 19

Graph showing the frequency of coded verbalizations over the course of the data collection period for Comprehension

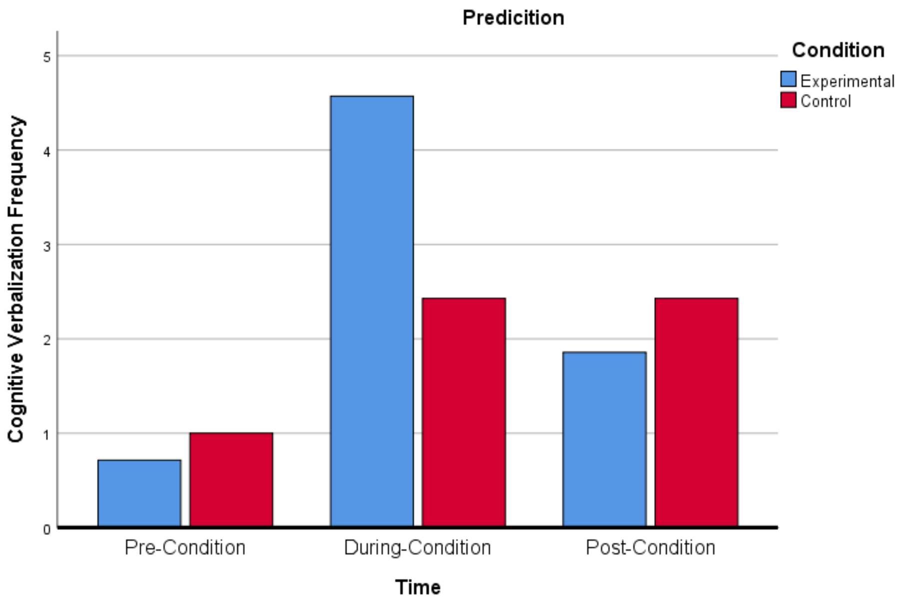


11.7.3 Situational Awareness - Comparison of *Prediction*

The differences in cognitive verbalizations for *prediction* from the pre-condition to post-condition measures showed an increase in the experimental condition (0.71 to 1.9, +1.19) and in the control condition (1.0 to 2.4, +1.4). This meant all participants provided approximately 1 occurrence of cognitive verbalization for envisaging a subsequent event, but this increased to approximately 2 occurrences of cognitive verbalization in the post-condition measures. The intervention materials were both able to increase the number of verbalizations that demonstrated a higher level of cognition in the simulation (e.g., '*the scrub nurse will have to ask the runner in 5 minutes for more swabs as she's running out of them*'). This suggested the participants were somewhat able to collate information further to form more consideration for the probability of an event, compared to the pre-condition. However, as both conditions had the same approximate increase it could be suggested the experimental condition did not have more properties which allowed augmentation of prediction compared to the control material (see Figure 20).

Figure 20

Graph showing the mean frequency of coded verbalizations per participant over the course of the data collection period for prediction. Both conditions increased verbalizations at approximately the same rate.

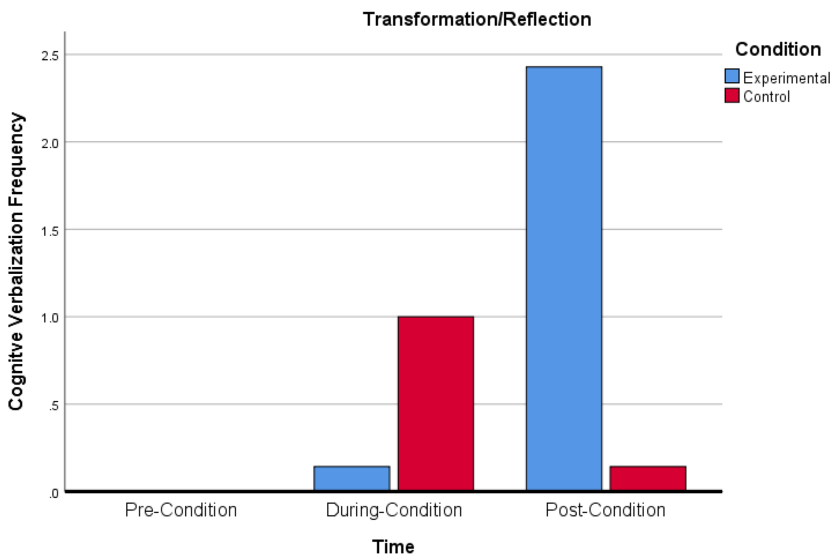


11.7.4 Situational Awareness - Comparison of *Transformation/Reflection*

The differences in cognitive verbalizations for *transformation/reflection* from the pre-condition to post-condition measures showed an increase in the experimental condition (0 to 2.43) and in the control condition (0 to 0.1) (see Figure 21). As participants had not viewed any material at the start of the experiment, they could not have had cognition-based verbalizations of reflection or transformation of data. An example of acknowledging more meaning from the same piece(s) of information was '*I didn't realize all the noise was affecting the surgeon when he is asking her a question, but I can tell now how much of an effect it has'*. The increase after intervention for the experimental condition would suggest an increase of reflection or transformation of cues- to develop better understanding of cue meaning(s). However, the control condition material enabled participants to do this also- this was not predicted as the material did not have specific cues within it. Participants may have applied the training material to reflect on past experience.

Figure 21

Graph showing the mean frequency of coded verbalizations per participant over the course of the data collection period for transformation/reflection.



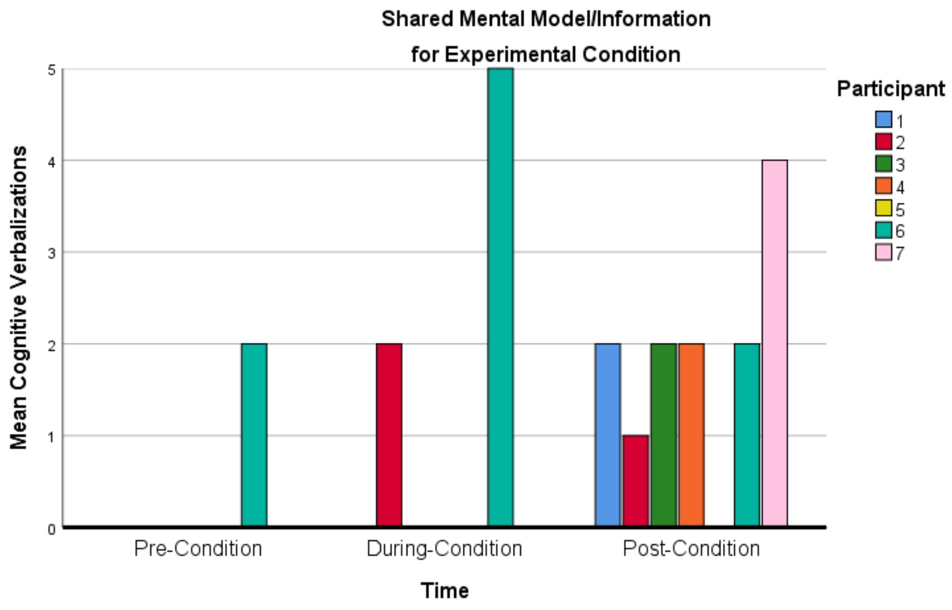
11.7.5 Teamwork and Communication - Comparison of *Shared Mental Models/Information*

The differences in cognitive verbalizations for *shared mental models/information* from the pre-condition to post-condition measures showed an increase in the experimental condition (0.3 to 1.9) but a decrease in the control condition (1 to 0.6) (see Figure 21). This meant that only the experimental condition was able to increase participants' abilities to verbalize their understanding that theatre staff had a concurrent set of perceptions and actions they attempt to maintain. An example may be- '*the scrub nurse knew to prepare it for the surgeon to change that, and he knew he didn't need to ask her to prepare it'.*

Participants 1, 3, 4, and 7 increased their verbalizations from 0 to at least 2 cues which suggested the experimental intervention material may have assisted participants to subsequently identify this non-technical ability in a realistic scenario. There were no verbalizations of this NTS for most participants during interaction with the intervention material, this was not expected as the intervention material had at least 1 event of this within it. However, this could highlight how the feedback received at the end of the intervention material guided their abilities to then identify the events present in the post-condition measure.

Figure 22

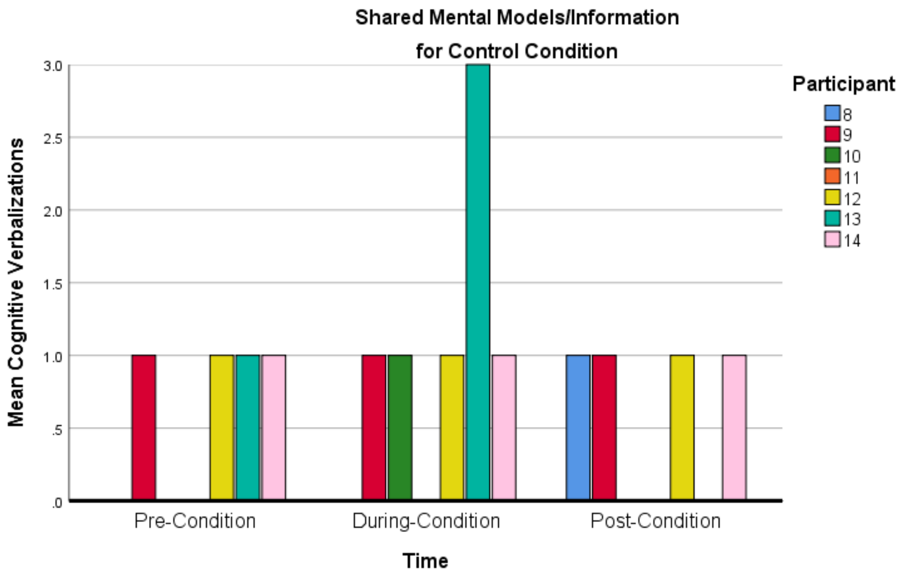
Graph showing the frequency of cognitive verbalizations for the experimental condition participants. Five participants had after intervention cognitive verbalizations which had reference to changes in shared mental models/information.



The control condition material explained and described this NTS and participants verbalized this once on average during intervention. However, no general improvement in identification could be suggested from the cognitive verbalizations as the frequencies stayed the same in the post-condition measure. The control condition material may not have facilitated participants to subsequently identify its occurrences when presented with a realistic scenario.

Figure 23

Graph showing the frequency of cognitive verbalizations for the control condition participants. There were mostly 1 occurrence in the pre-condition measure, and during the experimental material presentation which suggested little change.



11.7.6 Situational Awareness Descriptive Summary

The three levels of situational awareness from Endsley's (2000) model were explored by comparing the within-groups changes for the pre-, during, and post- condition measures. A fourth level was also explored. A notable occurrence from the data presented (notably when viewing the graphs) was the high increases in occurrences for the 3 situational awareness levels in the experimental group *during* exposure to their condition material. In the graphs this was indicated by the highest bars present. This suggested the material in the experimental condition had improved mechanisms to stimulate different cognitive verbalizations more than the control material. However, the critical goal was to explore if the exposure to both conditions could change cognitive abilities in the post-condition measure which had mixed results. The control condition material helped to increase the participants' cognitive verbalizations at the perceptual and predictive depth of SA- by approximately 1 cue each. The experimental material instead reduced occurrences of perceptual verbalizations by approximately 1, and had no effect on the frequency of *comprehension*, and showed similarity to the control condition as *prediction* also increased by a mean of 1. Transformation/prediction had a greater effect of change- being the reflection or reuse of the same information from previous performance to further understanding the simulated environment.

11.7.7 Inferential Findings for All NTS Categories

To clarify the mixed results described for the differences from the pre-to-post condition measure, inferential analysis was performed on the overarching NTS categories before being further broken down to investigate differences in each codings (See Appendix Figure 14A). A two-way univariate repeated measures ANOVA was the most appropriate test for analysis of the 4 main NTS categories as this test allowed investigation of the single within-subjects measures of time (before and after condition), along with between subject's independent measure of condition (being control or experimental). For this however, the dependent variable must be approximately normally distributed for all groups in the independent variables. The Shapiro-Wilk test of normality was selected as the sample size (14) dictated this was more appropriate over the Kolmogorov- Smirnov (Zimmerman, 2003). There were some significant deviations from normality for some of the data however more than 68% of the data was normally distributed with ANOVA being a robust test against the normality assumptions, hence tolerating the percentage of data that was not normally distributed. As a precaution however, it was found the NTS category of Leadership had significantly distributed data therefore separately analysed using a non-parametric Friedman's Test (See Appendix Figure 13). Homogeneity of variances must also be present for parametric testing. Therefore, this assumption was tested by Levene's test for Homogeneity of variances (Howell, 2011). Variances were assumed to be equal. The two way ANOVA was subsequently performed on the NTS categories comparing the frequency of cognitive verbalizations.

The results showed that situational awareness encoded verbalizations for the experimental condition before intervention (m=8.43, sd=3.46) did not significantly increase after intervention (m=9.71, sd=3.95) when compared to the control condition before (m=8.86, sd=4.22) and after intervention (m=11.57, sd= 5.48) (F(1)=.24, p=0.634). Decision-making also did not have statistically significant increases in respective cognitive verbalizations for the experimental condition before intervention (m=.57, sd=.79) did not significantly increase after intervention (m=1, sd=1) when compared to the control condition before (m=1.57, sd=1.72) and after intervention (m=2.71, sd= 3.45) (F(1)=.586, p=0.459). However, the results showed that Teamwork and Communication encoded verbalizations for the experimental condition before intervention (m=1.14, sd=.9) did have a statistically significant increase after intervention (m=3, sd=1.16) when compared to the control condition before (m=1.43, sd=.977) and after intervention (m=1, sd=1.27) (F(1)=6.26, p=0.028).

A Friedman's Test was performed on the Leadership based encoded verbalizations as it is the non-parametric alternative to the one-way repeated measures ANOVA when one group is measures at 3 or more points (i.e., pre-, during, and post- measures). However, as Friedman's Test ranks with blocks the data are allocated to either 1 or 2 and are then compared for positive and negative pair differences. This design would make the test entirely equivalent to a two-tailed sign test if data were measures at 2 time point. This meant between groups effects were not investigated however by referring to the means from the results of the within groups results this provided a clear indication of changes. The results suggested that for Leadership based encoded verbalizations there were no significant changes for the experimental condition before intervention (m=.14, sd=.38) compared to after intervention (m=.43, sd=.54) (Z= 1, p=.317). The changes for the control condition before (m= 1, sd=.81) and after intervention (m=.43, sd=.49) also were not significant (Z= 1, p=.83).

11.7.8 Inferential Results of Specific Coding Scheme Codes

As these categories were compressed into their respective NTS, this prevented the depth of analysis to explore the specific codes used in the coding scheme. There were approximately 5 codes which comprised an NTS category. Therefore, a secondary analysis performed further pinpointed the specific NTS codings(s) which produced the significant increase in the Teamwork and Communication grouping. Therefore, 19 codings (21 minus those for Leadership) within the Coding Scheme table were analysed using a two ANOVA, the coding scores were the dependent variable continuous data, IV one had two points of measurement (before and after), with condition (control/experimental) being the second independent variable.

The results showed that 3 of the 19 had statistically significant findings. The code SSI1, which was within the Teamwork and Communication category as named '*Shared Mental Models'* displayed a significant increase in frequency in the experimental condition before intervention (m=.29, sd=.76) to after intervention (m=1.86, sd=1.22) when compared to the change in frequency in the control condition before intervention (m=.57, sd=.54) to after intervention (m=.57, sd=.54) ($F(1)= 7.56$, $p=0.018$, $\eta p =.387$). This lower Fisher F value suggested the ratio of between-groups variance was only 2.37 times the size of the within-groups variance. The p value, being a function of F, was not significant as this represented the output of explained vs not explained variance not being able to conclude the groups were different enough at a population level.

Additionally, the results showed that the code SCS1, which was within the Teamwork and Communication category as named '*Taking Perspective of Others'* showed a significant increase. When reviewing the results, it uncovered that there were in fact no occurrences in any participants in the pre-condition measure, but 7 occurrences (across 5:1 experimental/control participants) in the post-condition measure. The change in frequency in the experimental condition before intervention (m=0, sd=0) to after intervention (m=1, sd=.57) when compared to the change in frequency in the control condition before intervention (m=0, sd=0) to after intervention (m=.14, sd=.38) ($F(1)= 10.8$, $p=0.007$, $\eta p =.474$).
Lastly, from breakdown of the 4 NTS categories, there was a significant finding that was previously merged with the other Situational Awareness codes which form that category.
The results showed that the code CT1, as named *transformation/reflection* showed a significant increase. As this was an NTS that can occur from the conversion of meaning of initially presented data, there were also no occurrences before intervention for all participants. It was after intervention that this ability may have been influenced to some extend from each conditional material.
This was from the change in frequency in the experimental condition before intervention (m=0, sd=0) to after intervention (m=2.43, sd=1.5) when compared to the change in frequency in the control condition before intervention (m=0, sd=0) to after intervention (m=.14, sd=.38) ($F(1)= 15.06$, $p=0.02$, $\eta p =.557$).

11.7.9 Summary of Findings

The key findings from the inferential exploration were as followed:

- Teamwork and Communication encoded verbalizations for the experimental condition before intervention did have a statistically significant increase after intervention when compared to the control condition before and after intervention.

- These were pinpointed to the themes of *Shared Mental Models and Taking Perspective of Others,* suggesting the experimental condition had significant effect compared to the control condition upon increasing the participants' frequency of these cues in identifying and understanding patient safety linked cues.

- The theme of *transformation/reflection* was also significantly increased after intervention in the experimental condition compared to the control condition, suggesting participants were provided with initial mechanism to attempt gaining new information from a cue by interpreting it in a novel perspective.

- There were no NTS related coding that were significantly different in the control condition over the experimental condition. This meant there were factors within the experimental condition which promoted at least equal output of cognitive verbalizations that allowed capture of the NTS of each participant

Within both conditions, the usefulness of investigating the array of NTS training material has benefits if the results are that researchers can better capture participant's NTS strategies. In doing so, improved assessment can be created with better targeting towards skills which do not meet a desired criterion. The usefulness of these findings was that the experimental condition material may better aid the extraction and standardization of the NTS abilities that were significantly increased by participants. However, these findings of increased verbalizations have one key limitation when referring to the goal in this thesis- an increase does not mean improved understanding towards those of a CAP. Therefore, the next step was to analyse participants' explicit cognitive occurrences before, during, and after intervention to assert *how* and *why* there were improvements due to condition over time.

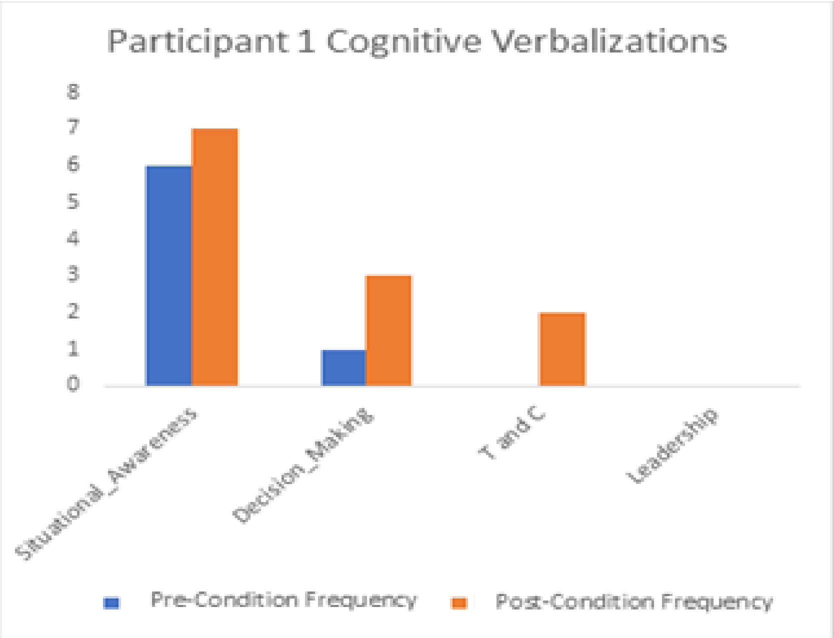
11.8 Before Intervention Measure- Cognitive Occurrences

Participant 1- Out of 7 situational awareness verbalizations 6 were specifically categorized into *Comprehension.* This category has been explained to have more cognitive complexity than simple perception to identify objects/actions, which does not describe how they link to each other. Participant 1 had an ability to understand the interactions of elements within the environment, and comprehended when these were or were not occurring. Within these, participant 1 showed rule-based expectancy for such interactions in the OT. For example, she stated '…*you shouldn't wipe a knife, it should be in a dish'*(I4). These phrases were formed due to an initial understanding of the expectations in the environment, but then the deployment of attention to a stimulus due to deviations of these expectations. This mismatch of understanding compared to the actual events affected the participant enough to speak aloud of each discrepancy.

However, there was only 1 coding result that fell within the *prediction* category- this meant participant 1 did not elaborate on these points to explain *how* these discrepancies may impact the environment in the short- or long-term future. There were no verbalizations which were coded into either teamwork and communication, or leadership which also suggested the participant had less focus on the social cues and events (i.e. than between 2 or more individuals). Indeed, 6 items in the theatre were mentioned. Additionally, words or phrases such as '*attention*' (I1), '*glancing around'* (I6), and '*is someone behind me'* (I7) indicated they were more focussed on the cognition-based categories of the NTS taxonomy. Therefore, specific cues and events between an individual and equipment were more common. Figure 21 shows the frequency of coding for the before and after measures.

Figure 24

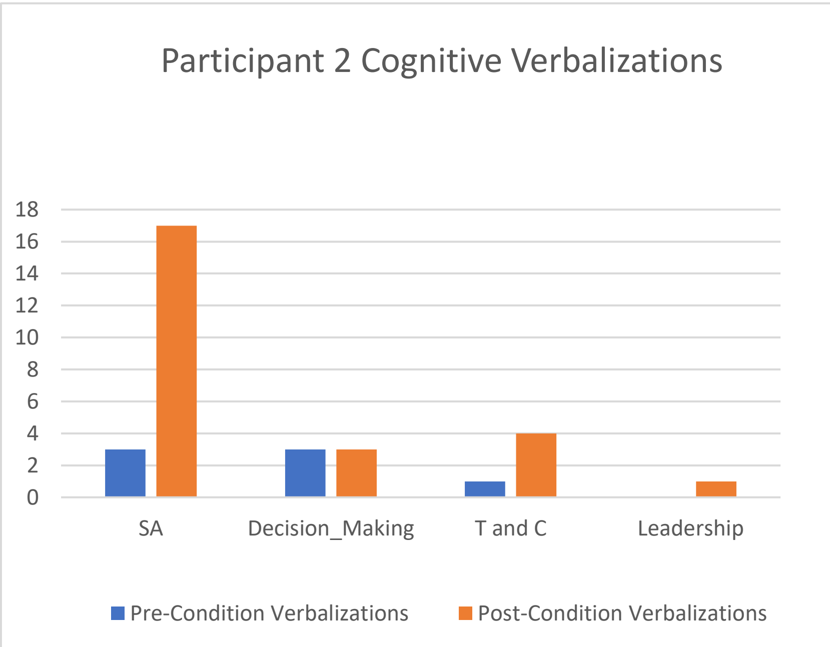
Graph showing the frequency of cognitive verbalizations for each non-technical category. No leadership orientated reasoning was verbalized but increased frequency occurred for the other categories.



Participant 2- Overall there were 18 less encodings before intervention measure compared to after intervention. This may have been due to the manner in which the participant provided their thoughts and may have been a sign of cognitive overload from the novel setup. She stated '*So I didn't really speak during the video but just pausing at the end yeah I should have paused it.'*. Furthermore, there was a suggestion of attentional bias that the participant recognised. This had a negative impact on their ability to look around the environment, but it again demonstrated self-awareness which was a higher-level skill. She stated, '*I know there were several things going on around me, but I didn't really look as I was anticipating something surrounding the surgeons*.'(l6). This was an unpredicted finding however demonstrated that a wide array of skills and strategies may exist within the same cohort. There were frequent identifications of events occurring at a basic perceptual level. Most of the coding fell within the *perception* of the staff in the scenario, and suggestions of *information gathering* towards making decision. This would concur with the previous decision-making measure in that this participant has a developed understanding of what areas, cues, or information should be attended to strengthen understanding towards better decision-making. Figure 22 shows the frequency of codings for the before and after measures.

Figure 25

Graph showing the frequency of cognitive verbalizations for each non-technical category. There was incongruity with the increase in SA compared to the other 3 NTS categories.



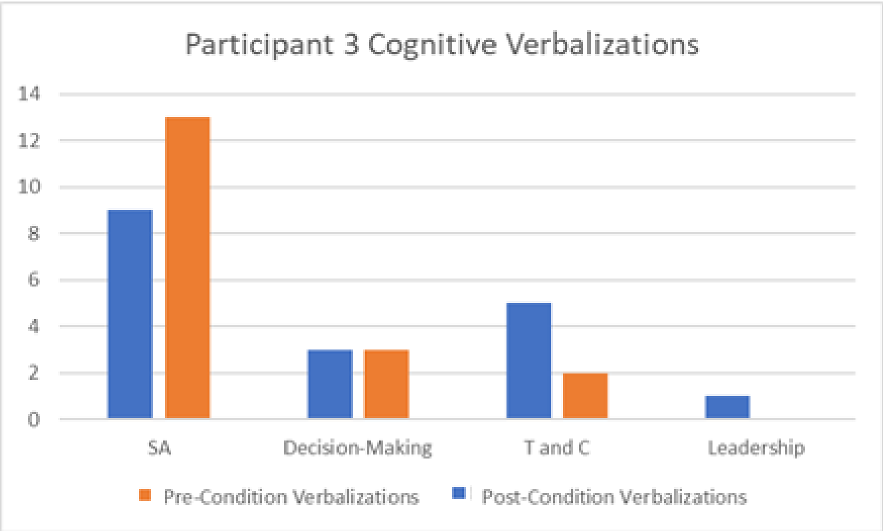
Participant 3- Participant 3 was randomly selected for the inter-rater reliability check for the second independent coder. The codings of verbalizations from participant 3 suggested that a substantial amount of knowledge and cognitive processes fell within the first level of situational awareness- being *perception.* Eleven out of the 17 occurrences (65%) were analysed to posit a basic perceptual acknowledgement of either the action of an OT member or use of tools. These were categorized in *perception* as the model defined by Endsley

(1995) described that level 2 situational awareness requires a synthesis of the previous elements in level 1, in order to comprehend the situation. The information of basic tools, individuals, and actions should be transformed and integrated by the individual's cognitive development. This is in order to understand the role and significance of these elements. The cognitive walkthrough from participant 3 suggested they were able to recognize situation related cues but did not expand on the meanings, consequences, or rationale behind the occurrence of these events. A few examples were, '*Somebody's writing on the board about what's going on'* (I3), '*And they're opening a new pack of somethin*g' (I6), '*Looks like she's also counting and then writing it up on there.*'(I7), and '*something dropped down there as well. And there.*'(I14).

Because of the participant's restrictions in comprehensive (level 2) and predictive (level 3) explanations, they were unable to show they could compile the effects from interactions which may have been defined as 'minor' undesired events. These could compile to become major undesired events as explained by the Swiss Cheese Model (Reason, 1990). This was also verbalised by the participant as they stated 'I *don't really think I've seen a risk of an error yet.*'(I8), which was a prime example of how Verbal Protocol Analysis and coding in the defined format had the ability to explore the tacit non-technical techniques and strategies, which may then be supported by the participant's explicit behavioural events. Overall, participant 3 showed a reduced amount of higher level NTS abilities when compared with the majority of participants, but was able to identify many basic events that were related to error even if they did not elaborate why. Figure 23 shows the frequency of codings for the before and after measures.

Figure 26

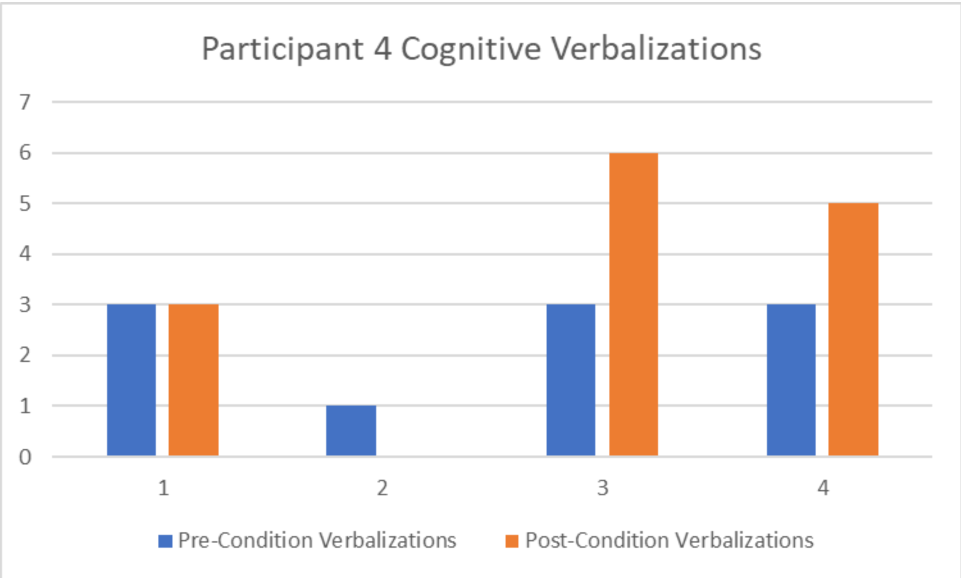
Graph showing the frequency of cognitive verbalizations for each non-technical category. The participant had many perceptual situational awareness verbalizations that did not include further explanations. Teamwork and Communication, and Leadership had marginal increased verbalizations.



Participant 4- Participant 4 referenced how she interpreted the scenario and its effects on event understanding. This helped her to link the effects of cues in the environment. For example, she stated '*I'm feeling aware of the noise, I didn't notice the music either, I don't know if it's on all the time or not but with the talking then everything does add up.' (I1-3).* This self-awareness was then more clearly verbalized when she stated, *'I've kind of pieced together the noise is making the communication harder'*(I6). This self-awareness also acted as a tool to understand any assumptions, she stated '*nothing happening behind me I just assumed there wasn't anything before I'm usually just looking at the scrub practitioner area'(I4).* The reflection suggested a conscious effect to understand how and why she was identifying cues and could suggest a stronger NTS ability. A limitation was that when paired with her ranking scenario analysis there were no encodings of level 3 situation awareness therefore the participant did not demonstrate the ability to combine cues to suggest events that could occur in a short time frame. Figure 24 shows the frequency of codings for the before and after measures.

Figure 27

Graph showing the frequency of cognitive verbalizations for each non-technical category. The participant had many perceptual situational awareness verbalizations that had little explanation. Teamwork and Communication, and Leadership had decreased verbalizations, which was not expected.

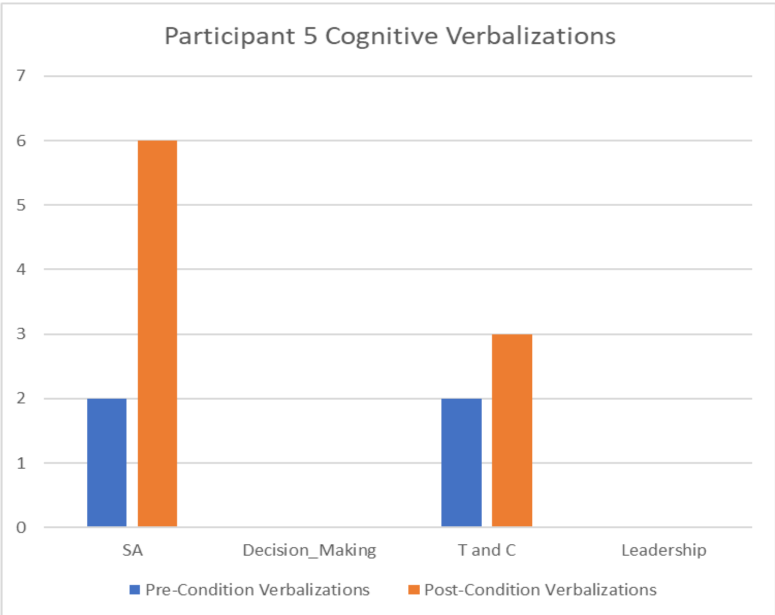


Participant 5- Participant 5 was the only individual who did not expand upon their rationale for the rankings in the decision-making scenario. The absence of any decision-making encodings in the pre- and post-conditions may be linked to this, and suggested this participant had an unknown problem with eliciting their reasoning or that they did not have the strategies to be able to expand upon. There were minimal verbalizations with participant 5. When paired with the lack of verbalizations for the decision-making measure there was suggestion, she may have had a barrier that prevented these verbalizations from occurring. The researcher added a note in the transcript that read *'was able to talk when not doing the videos, but just not talking much during the videos*'. There may have been considerable difficulty for participant 5 in unpacking her cognitions in a real-time manner, instead better at talking afterwards. This was similar to participant 2 who also stated she wanted to talk after the scenario rather than in real-time. These preferences/abilities of verbalizations were not found in previously reviewed research and were not accommodated for in this thesis.

From the transcription there were just 4 NTS categories identified. There were two codings that fell into communication- specifically *exchanging information*. Two codings were categorized into comprehension- specifically *increasing/decreasing others SA.* The participant demonstrated they were able to change their perspective to assess the effect of reduced SA. I.e., *'the circulating staff aren't paying attention, they are talking in a group and I know they don't have to do anything until someone asks but they are not being involved* (I3)*'.* The depth of exploration for participant 5 was limited for measures before intervention and was considered in post condition measured to avoid unfounded assumptions of change from occurring. Figure 25 shows the frequency of codings for the before and after measures.

Figure 28

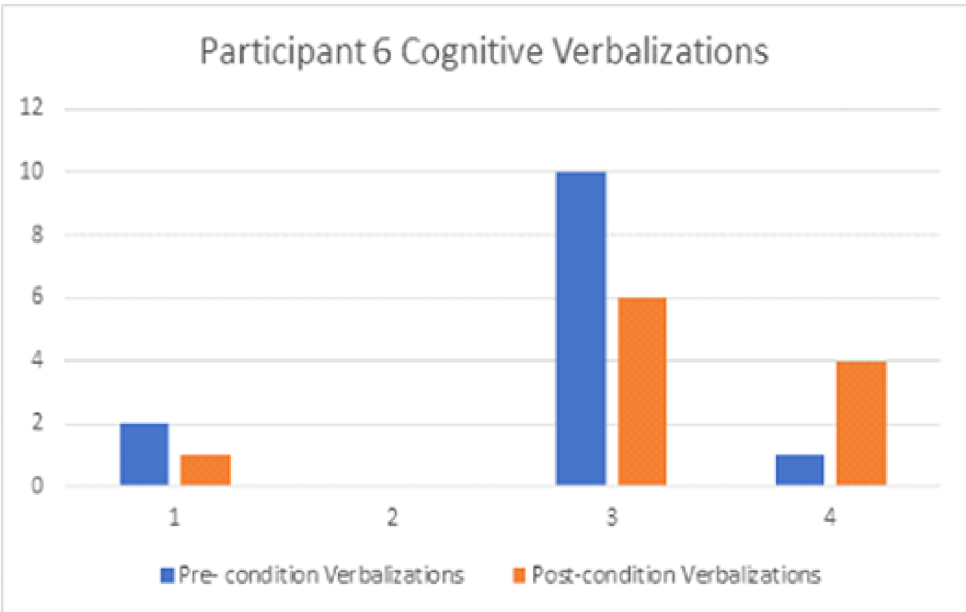
The NTS expressed by the participant only had components of two of the four categories. The independent second-rater also matched with this coding outcome.



Participant 6- Participant 6 was randomly selected for the inter-rater reliability check for the second independent coder. There were 14 non-technical related cognitive verbalizations coded overall, the majority were in the *process deviation* and *reducing/increasing SA* codes in the situational awareness category of *comprehension*. Indeed, they had 3 and 5 occurrences respectively. A large portion of the transcript was highlighted with categorization into *reducing/increasing SA,* as the participant focused their explanations around this more than other elements. This participant did not identify tool or process related events and cue, but awareness-based cues. This meant that approximately 60% of the transcript were sentences capturing the thought processes in relation to *reducing/increasing SA.* Examples included, '*getting a bit distracted there'*(I2)*, 'ODP is still keeping an eye on the procedure whilst she's talking'*(I7)*, 'both looking and making sure they're still on the ball'(I8),* and *'she's positioned herself in a way that she can't see what's going on around her (9)'.* However, this implied that the participant's ability to identify other non-technical related cues was not established due to the majority of verbalization towards the same incidents. Figure 26 shows the frequency of coding for the before and after measures.

Figure 29

Graph showing the frequency of cognitive verbalizations for each non-technical category. The participant had many perceptual situational awareness verbalizations that had little explanation. Teamwork and Communication, and Leadership had decreased verbalizations, which was not expected.



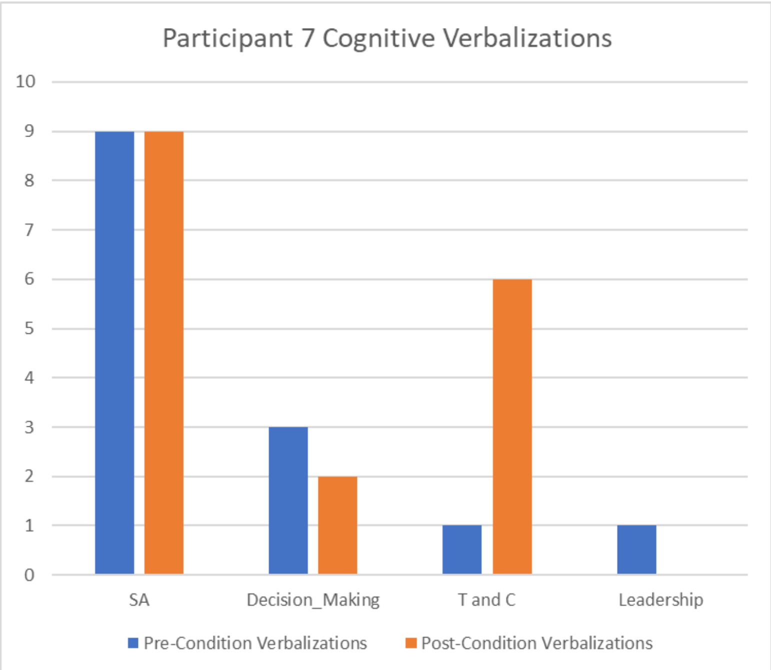
Participant 7- The previous prediction - that there should be more occurrences of complex cognitive verbalization and less basic perceptual, or rule-based strategies- was not correct. Eight of the 16 (50%) encoded occurrences fell within the perception and information gathering categories. This suggested participant 7 deployed attention towards an initial scan of the environment before employing further advanced non-technical evaluations. This has also been evident with other participants' identifications- i.e. participants 2, 4, and 5 had initial information gathering periods. Examples from participant 7 were:

'*The anaesthetists there talking and the monitor. The surgery is a head surgery but I can't see what it is'(I2), 'So there are quite a few people in here I'm guessing it's a complex surgery 'cos there are 2 anaesthetists and 3 runners helping'(I3),* and *'The surgeons asked someone over, he's making her look at something'(I6).*

Therefore, although participant 7 demonstrated critical decision-making abilities that had developed above rule-based reasoning, in the before intervention measure they deployed attention to gather cues and information within the environment. There were few elaborations of identified cues, with only one *process deviation* related process coded. This referred to a trainee being close to the sterile field, and consequence that '*she might make it not sterile'*(I6) and was the most commonly occurring cue observed by all participants. Figure 27 shows the frequency of codings for the before and after measures.

Figure 30

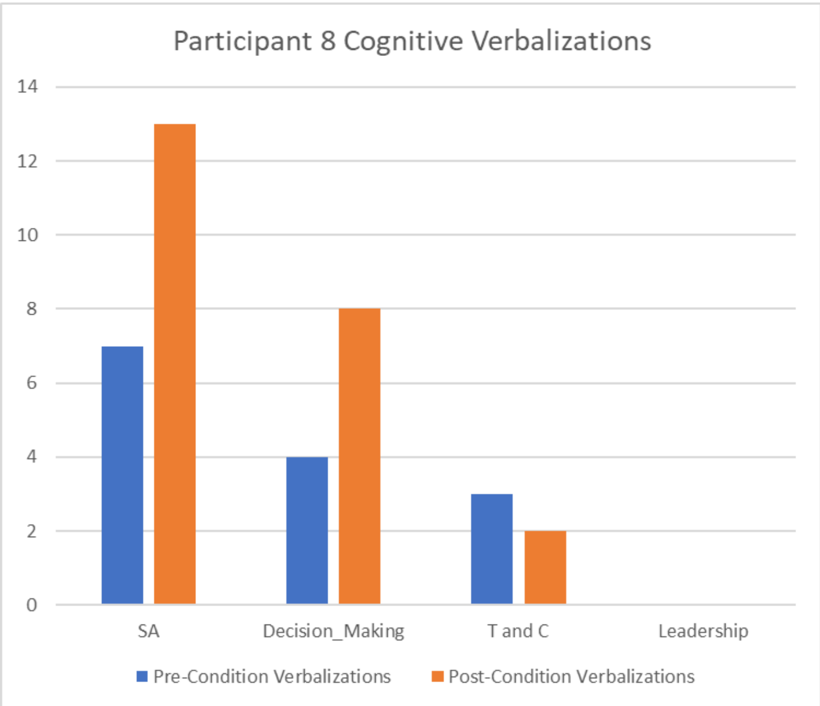
Graph showing the frequency of cognitive verbalizations for each non-technical category. A pertinent amount of NTS were relating to situational awareness before intervention and this level was maintained after intervention



Participant 8 - Participant 8 was randomly selected for the inter-rater reliability check for the second independent coder. They had applied previous experiences to emphasize their point on the cues initially identified. Although this may simply be the method to which the cues were expanded upon, their experience may have been a factor in their initial scanning of the environment as they have first-hand experience of the consequences of these issues. The participant detected cues relating to operating theatre overcrowding and increased risk of desterilization. For example, she indicated more than once that an individual who was not sterile was too close to sterile items. This was a common occurrence in most of the participants. The deployment of attention towards these cues meant several verbalizations were coded in the *perceptual* and *comprehension* cognitive categories and were orientated around procedural surgical events more than the social non-technical skill categories. Although she identified cues relating to distractions towards the end of the scenario, she had not identified teamwork-based equipment tracking issues. Instead effort in both identifying and explaining cues was directed towards congestion and increased risk of desterilization. This may have been a higher priory for the participant and highlighted that different experiences may alter the way participants experience and explain different environments. Figure 28 shows the frequency of codings for the before and after measures.

Figure 31

Graph showing the frequency of cognitive verbalizations for each non-technical category. The scenario had few leadership cues, however like participants in the experimental condition there were no coded leadership verbalizations.

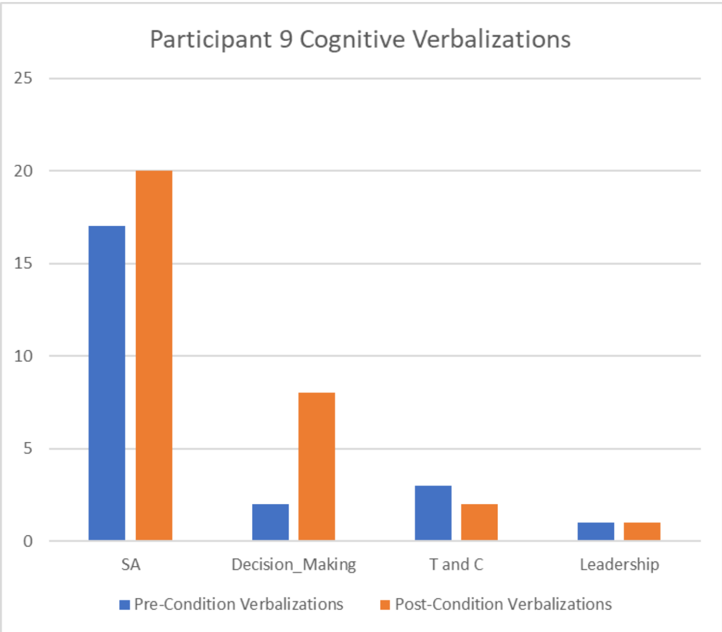


Participant 9- Participant 9 was randomly selected for the inter-rater reliability check for the second independent coder. The pattern was similar to most participants in both conditions as there were more cognition based codings than the socially orientated codings. Out of the 17 situational awareness cognitive verbalizations encodings, 7 were articulated for *process deviation.* As concurred by the influences in their decision-making in the previous paragraph, these verbalizations were regarding the identification of behavioural processes of staff members based on the standards of practice- i.e. following the rules of infection control. She stated '*when you're in the theatre, you have a square which is like the surgical field, you're not supposed to be as close to the patient as this lady is'* (I16). Indeed, most of the cognitive verbalizations were cues at a larger visual level than compared to smaller equipment cues (i.e. a small tool in someone's hand). This demonstrated level 2 situational awareness skills as there were frequent identifications of the dynamic interactions occurring from the staff.

In terms of prediction, a large number of verbalizations elaborated on the meaning of only one cue. For example, a single equipment related cue was cognitively assessed by the participant for its consequence of being a trip hazard. However, a small portion of verbalizations also described a chain of cues towards a later event, which could indicate some developed predictive ability from the combination of 1 or more cues. These cues could be unrelated to patient safety if each cue was viewed independently. She stated, '*So what the scrub practitioner's doing there is leaning on the bed, so that could be risky...the patient could move so that could be risky for the surgeon'(I5).* Although an uncomplicated observation, the understanding that minor and momentary event may cause a chain of events demonstrated practical application of the Swiss Cheese Model (Reason, 1995) in that minor cues should be identified and prevented to minimise the combination of their negative effects. Figure 29 shows the frequency of codings for the before and after measures.

Figure 32

Graph showing the frequency of cognitive verbalizations for each non-technical category. A pattern similar to the majority of participants



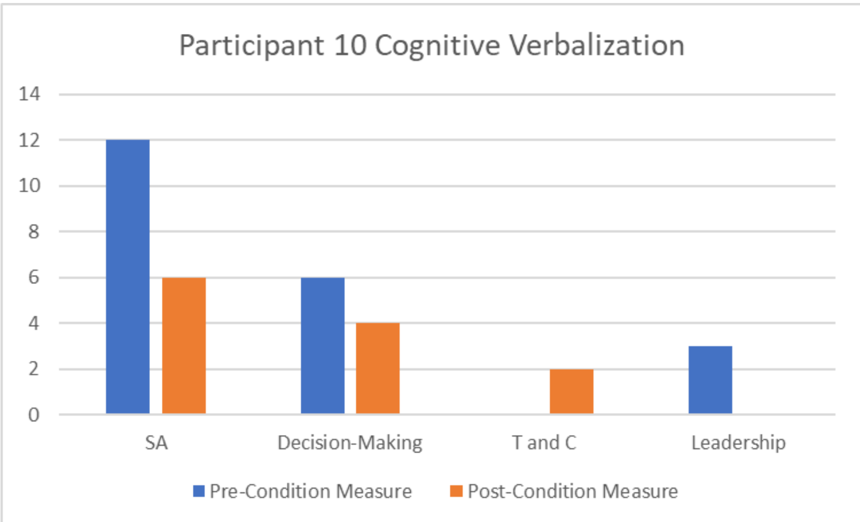
.

Participant 10- There was an array of cognitive verbalizations which delineated why specific cues were mentioned. The participant demonstrated a pattern that identified a cue then followed it up by more information from a safety viewpoint. This was consistent throughout their transcript. This meant the cognitive events during scanning of their environment were explicated within the level of *comprehension* as she was able to explain how the events deviated. Such examples were '*The scrub nurse hands over the blade to the surgeon using her hand, ideally, she should be handing it over in a kidney dish'(l13).* Most expansion was regarding procedural consequences, but there were verbalizations which detailed the effects on the patient.

In terms of prediction, the participant showed two basic predictive cognitions when she stated '*She could fall over, she could possibly fall onto the patient'*. This prediction had a short temporal window and did not have many embedded or complex cue interactions. Indeed, many identified consequences were described only within several seconds and were general statements. For example, '*there's the suction pipe behind the lady and she attempts to hop over it. Again, that could be a trip hazard for anyone'* (l19). This did not mean they were of less importance for patient safety, but the participant's predictive ability as defined by Endsley's (2000) model would be at its earlier and less specific predictive stages. Based on this, it could be suggested the participant's NTS can produce a range of non-technical event types, but may have a limited capacity to combine cues and a limited detail of events predicted. Figure 30 shows the frequency of codings for the before and after measures.

Figure 33

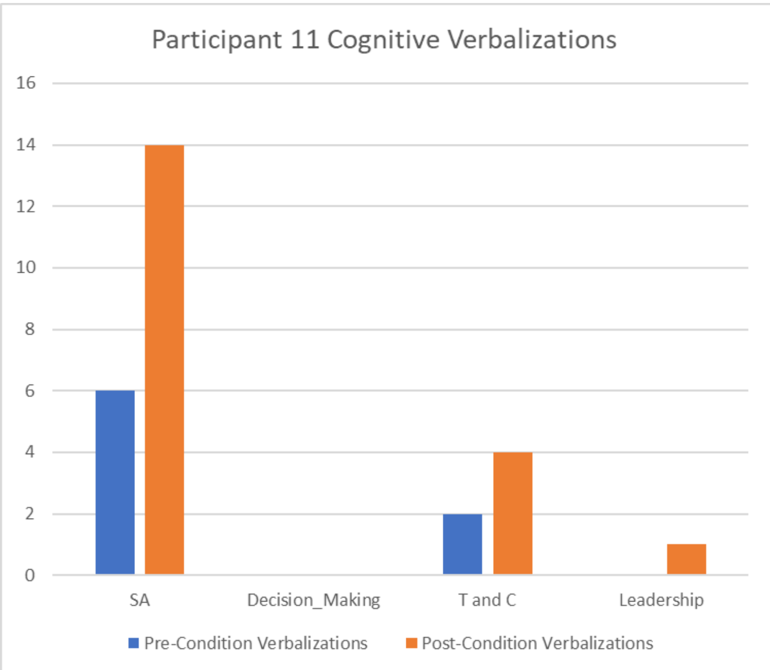
Graph showing the frequency of cognitive verbalizations for each non-technical category. There were less verbalizations in the post-condition measure which may have been the contributing limitation in verbalization. The reason for this was not apparent.



Participant 11- The participant's decision-making was not present in any of the three conditions even though they performed in alignment with the 57% of participants in the ranking scenario and verbalized adequately. This only occurred for participant 11. Instead, the majority of encodings fell within situational awareness for both before- and after intervention conditions. Participant 11's data was checked twice with the longest time point possible between the two checks to counter any intra-rater changes that may have occurred. Six out of the 8 (75%) occurrences were categorized in the cognitive process of *comprehension.* Four of these were surrounding the reduction of situational awareness. There was an analytical emphasis of verbalization explaining how signs of miscommunication were caused by reduced situational awareness of the staff (e.g., '*and once the surgeon or anaesthetist says something, they have to say pardon which means that they weren't paying attention*'-*I3).* These descriptions of *comprehension* and awareness account for most of the cognitive verbalizations. These are critical as Mushtaq et al (2018) suggested communication issues contributed 16% towards surgical incidents. Improvements in awareness of these cues may help to reduce such problems. However, this meant many other cue types were not present and that most of the coding scheme categories were absent. There were no occurrences of *prediction,* or any suggestion towards the participant's uses of *decision-making* in the scenario which limited the ability to explore further than the codings around the level 2 situational awareness verbalizations. Figure 31 shows the frequency of codings for the before and after measures.

Figure 34

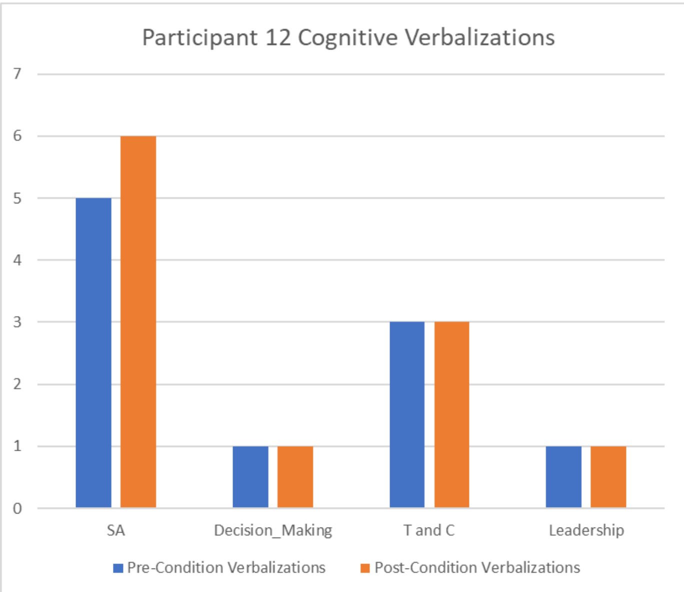
Graph showing the frequency of cognitive verbalizations for each non-technical category. The participant showed attentional preference towards the communication and awareness of others but other event/cue descriptions such as tool usage or workload were absent.



Participant 12- Participant 12 was inter-rater checked by random selection as described in the previous inter-rater result section. Being a typical cue identified, the positioning of staff was verbalized and a rule-based explanation was provided that they should keep a minimum distance between themselves and the patient. The participant perceived the different 'groups' of staff within the scenario and elaborated on the consequence of this being that they were not directing their attention to the patient. This comprehension was followed by the participant trying to make sense of the staff displaying reactive behaviour. She verbalized what she believed should be occurring, which demonstrated application of the 'barriers' within the Swiss Cheese Model in that individuals should be proactively implementing measures/behaviours to create a preventative barrier for undesired events. She stated, '*The staff around the side they need to be ready, just in case they need something or there's an emergency situation, they all have to be ready because they are a lot more relaxed having their own conversations* (I8)'. These responses concurred with the previous suggestion that participant 12 was patient-safety orientated as the underpinning influences were driven by error prevention. She had perceptual and comprehensive skills to identify such cues and could interpret their meaning in a more developed manner. Figure 32 shows the frequency of codings for the before and after measures.

Figure 35

Graph showing the frequency of cognitive verbalizations for each non-technical category. This was the most similar between the pre- and post-condition measures. The control condition material did not affect participant's frequency of cognitive verbalization in the scenarios.

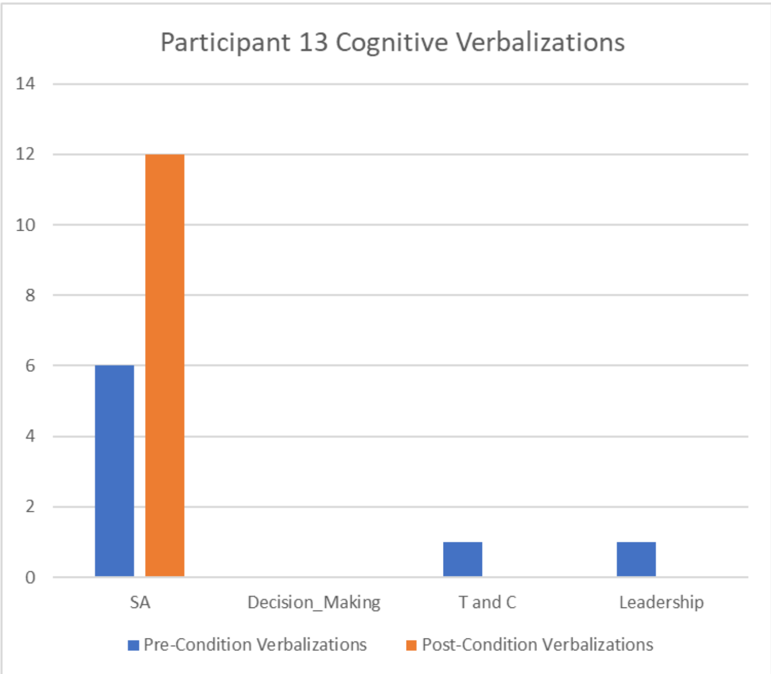


Participant 13- Lower amounts of coding verbalizations occurred in 3 of the 4 NTS conditions. Similar to participants 7, there were identifications of events but described them in isolation (i.e. '*there's lots of equipment on the patient* (I1)', '*she's coming over to assist* (I2)', 't*he trainee, she just steps over some wires* (I7)'. There were error prevention related descriptions of positive cues, e.g. '*It's really good that the surgeon is talking to the runner about what is going on, an involving all the people* (I4)'.

However, there was no further predictions of the effect of these cues and although the before intervention measure instructed to identify NTS or patient safety related cues/events it also instructed to explain the implications and protocols surrounding them. Participant 13 also did not include predictive cognitions or more in-depth deliberation within the ranking scenario. There were no codings that captured decision-making strategies or logic, even at a simpler rule-based level. The application of the Swiss Cheese Model (Reason, 2000) for this data would highlight the ability to identify independent minor issues but not to predict events if they were combined to create a developed understanding of their effects. Figure 33 shows the frequency of codings for the before and after measures.

Figure 36

Graph showing few verbalizations surround the protocols viewed in the before and after intervention measures. There were 6 encodings that were at the perception and comprehension level.

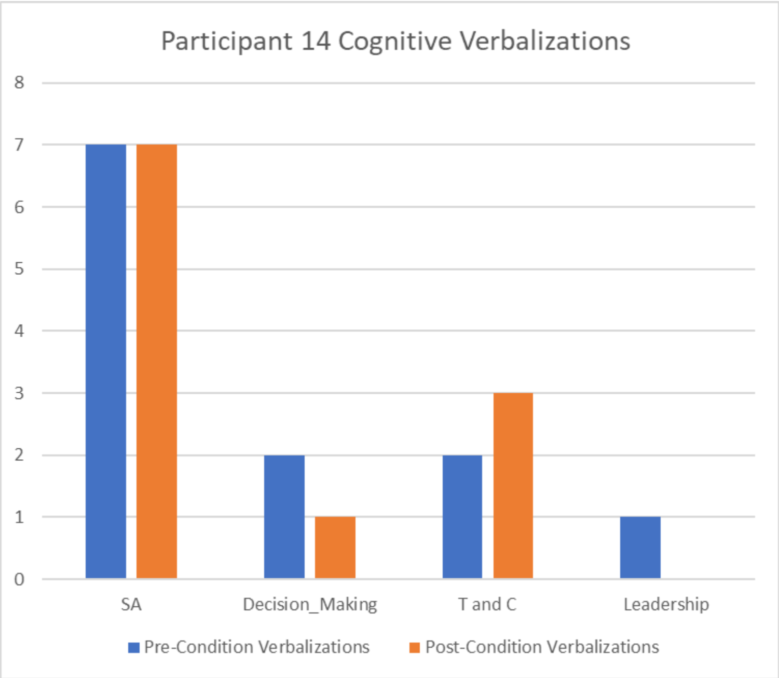


Participant 14- There was a mixture of behaviour related cue identifications, details about the surroundings, and team-based communication. The result was that the participant identified the cues that were also commonly identified by most other participants. For example, two popular SA level 1 and 2 cues were given: '*So, the scrub nurse passed the knife by hand… The observer just hopped over the suction there (I8)'*. However, a less common theme was the surgeon attempting to keep the team informed of the current tasks (shared mental models) of which the participant identified and elaborated upon. I.e. *'what's that the surgeons making the runner have a look I don't know why she's not a surgeon, but I suppose he's letting her learn about the procedure, she wasn't really interested but it's good to share*(I7)'.

These cognitive verbalizations were adequate for patient safety for the participant to believe they had found some evidence of potential issues but similar to participant 13 did not combine these cues. Participant 14 stated she did not '*see any major problem, just little bits and bobs of things (I10*)'. This phrase suggested she may not have been able to deliberate if these smaller events could lead to bigger issues as defined by the Swiss Cheese Model (Reason, 1995). Figure 36 shows the frequency of codings for the before and after measures.

Figure 37

Graph showing the frequency of cognitive verbalizations for each non-technical category. The same trend occurred with SA and TC acquiring the mist verbalizations.



11.8.1 Summary of Findings – Before Intervention

The exploration suggested all participants had level 1 perceptual skills, with most participants using level 2 comprehensive abilities (less evident in participants 5 and 13). Most did not use predictive power to identify potential undesired events- only participants 1, 9, 10, and 12 had predictive statements in their transcripts. Participants 1, 9, and 10 also had higher frequency of cognitive verbalizations towards identifying deviations of expected processes. There was partial mirror to the exploration of ranking orders as there was a range of strategies and depth of cognition, especially in terms of situational awareness. Most participants had focus on the type of cues relating to team processes. Fewer participants instead orientated around equipment and other cues not directly linked to the OT staff (participants 2,4, 8, 10, and 14).

The analysis before intervention explored each participant with reference to common non-technical and human factors derived skills from previous literature, from this variance in individual differences was further captured. Within the same participant there were alterations in how a scenario was understood, from the initial written decision-making ranking scenario to the surgical video viewed before intervention. This posits that the continuation of individual assessment during intervention would provide the same rich detail into their efficacy on participants. The measure after intervention could focus on the collective changes of one group compared to the other. Therefore, individual analysis of cognitive verbalizations aimed to explore the effects of the intervention material in developing participants' NTS abilities.

11.9 During Intervention- Cognitive Occurrences for Control Condition

As the control condition was the representative basis/benchmark of NTS training, the control condition participants (8-14) were firstly presented and summarized to generate the outcome of its effects on participants. This would help contrast the differences when the results from the experimental condition participants (1-7) were presented. The material in the control condition were a compilation of videos that represented the content of contemporary training courses that can be accessed online or with enrolment of workshops and other organised events. The material explained non-technical and human factors theory that is similar to material used in domains outside of healthcare.

Participant 8- The three levels of situational awareness based on Endsley (2000) were described and a contextual example was provided that suggested the participant listened to, interpreted, and explained it in their own manner. However, for level 3 *prediction* there was issue when the participant tried to deploy the knowledge gained from the control condition into a practical context. He stated, '*I'm not sure how I'd be able to predict things- like I can only predict what part of the surgery is next because I know the stages of the surgery' (I5).* The participant expanded to state '*to predict you need to have mental models but I'm not sure how, I mean unless I'm there maybe I will see something that makes me think something bad might happen' (I6).* This demonstrated that there was not enough context from theoretical information for the participant to reflect on their own non-technical skills. This was an initial indicator from a participant that could be a suggested difference between the 2 conditions in this project. Real-life context can allow an individual to apply their knowledge into a reflective practice that can help them adapt their abilities based on this reflective practice. For decision-making, the participant verbalized how the strategies were able to be interpreted as 'easy' to 'hard', he stated '*left to right that's kind of easy to hard'(I9).* The meaning of the more complex strategy of *creativity* was hard, and more immediate strategies were simpler in nature and 'easy' to perform. This concept may have been due to an image when the decision-making strategies were being explained (see Figure 29).

Figure 38

Four types of decision-making strategies as visually depicted in a non-technical skills lecture (From surgicalsciences, 2017).

This suggested the participant understood that each strategy has various components and verbalized his attempt to apply them to his own skills- he stated '*I get told not to touch things still if I'm not sterile, like basic rules. But over time you know more the expectations than the rules like if it's the last surgery and it's all over then, yeah you can touch anything as it's all done with…that's more analysing the situation.' (I12).* Lastly, a reflective coding occurrence was captured which showed meta-awareness based on the content of the control condition- '*It's good to know and if I'm doing something next time I'll be thinking, right what kind of decision making am I using now and is it the best'(I15).* The effects of this on the post-condition measure and SAGAT were subsequently analysed.

Participant 9- Participant 9 stated she did not know about the Swiss Cheese Model and talked aloud about the way she perceived the information may become useful in surgery. Yet, as described in the measure before intervention, she demonstrated practice of the Swiss Cheese Model. During the control condition presentation, she then verbalized her initial theoretical understanding of this. Participant 9 stated, '*you can kind of anticipate a situation, you can stop situations getting worse than it already can be by anticipating the situation and having awareness of your surroundings* (I7).'

The conceptual understanding of shared situational awareness was also evident in the latter part of the transcript as there was appreciation for the transfer of knowledge to enhance the teams understanding of the environment; she stated, *'Each of them has their own sort of knowledge but it's not about how much knowledge* *they have…it can be useful if you have some knowledge from a scrub nurse as well that could help an anaesthetist or a surgeon'(l20).* Lastly, from the control condition material the participant understood the role of leadership towards patient safety. Leadership was coded once in the pre-condition measure and was the identification of the surgeon directing staff (pre-condition transcript line 12). The theory behind it was verbalized further: *'I agree with the statement made by Mazocco as communication is theatre is very important…There should be one leader speaking in theatre, leading everyone in the right sort of direction and it obviously prevents more complications.'* Overall, the participant demonstrated appreciation for the theoretical information provided and was able to convert this into their interpretation of how the non-technical phenomena described would improve patient safety in theatre.

Participant 10- The participant verbalized how they processed the information gained from the control condition pertaining to the different decision-making strategies. She stated she did not know about them and articulated an analogy to understand how they would be applied in a certain context. She stated '*I didn't know about the ways people decide, like the strategies. But I suppose it's like learning to drive, you follow the rules of how to use the car and on the road but eventually you don't have to think about, and you can spot problems better' (*l3-5). This demonstrated she understood the concept and was able to apply the knowledge appropriately therefore had the information to relate this concept to her own situational awareness in theatre. In relation to the Swiss Cheese Model (Reason, 2000) the participant also verbalized her understanding of the model in that minor events can frequently occur if human elements allow them to bypass preventative measures. She stated '*it's the little mistakes that you might make, that eventually just may become a habit that could put any patient's life at risk'* (l10). Participant 10 did not verbalize her non-technical cognizance towards shared mental models and patient safety but these were the only skills from the control condition she did not refer to. For example, for situational awareness she applied the concept to some form of real-world meanings similar to participant 10- *'I work as a scrub nurse in theatre, so I believe having situational awareness is very important… I feel that I can be more aware of my surroundings, not just think about the patient's safety but think about my own safety* (l12-13).

Participant 11- The cognitive verbalizations were initially describing the effects of leadership on students' confidence levels, as the participant was expressing an example of how they believed differences in non-technical skill can affect colleagues' experiences and delivery of care. Then the participant acknowledged that the overarching goal for the OR team was patient safety with communication being their suggested most valuable skill. This may have been a reason why their pre-condition measure was primarily communication orientated. The 3 levels of situational awareness were expanded upon by applying it to previous experience (similar to participants 9 and 10)- she explained, '*some students they have a perception of, and obviously they understand that they have to stay out of, the sterile field if you're not scrubbed up. But then some members of staff think just because you're a student you're not aware* (l13)'. This application of the theory provided by the control condition material helped the participant to better make sense of previous events, which demonstrated reflective elements. She was immediately able to understand how the abstract information was relevant to her actions, she articulated '*ODP's job role is to…always to anticipate what the anaesthetist or the surgeon or the patient require and then by using this model by projecting and understanding what's required you're in a better*

(I18)'. This was strong evidence to suggest the control condition material had the potential to improve outcomes in the subsequent post-condition measure and modify the non-technical verbalizations produced without practical scenario interactions.

Participant 12- The participant acknowledged her understandings of the content of the video and paused frequently to provide thoughts. She also gave examples of how she would apply these NTS skills in real-life or how they can affect performance if not implemented correctly. This was consistent in participants 9, 10, 11, and 12 and a key finding on how individuals make sense of this material. There was an excellent interpretation of SA level 3 prediction, as the participant considered anticipation to be a proactive effort, whereas intuition was a more automatic feedback process. This was accurate as intuition has highly developed mental models of events that become automatic to implement (a 'gut-feeling') (Ericsson, Hoffman, & Kozbelt, 2018). Whereas anticipation is the conscious effort to identify cues that aid in understating future events. Generally, the participant demonstrated the understanding of the key concepts and patient safety meanings extracted from this content. An example of understanding communication with reference to the Swiss Cheese Model (Reason, 2000) was, *'Lack of communication could lead to a patient getting the wrong treatment, which could lead to injuries, potentially worse things* (I16)'. Understanding transformation or reflection of information she stated '*So situational awareness I think is linking onto a reflective practice. What would you do better next time? What have you seen? What have you done? What could you do to improve?* (I18)'.

Overall, the participant showed process of listening to, and converting the information into examples they could relate to. This process demonstrated how individuals in this control condition can try to make sense of conceptual models or theories by application to real world experiences.

Participant 13- There was a negative effect on the participant as they learned about the ranges and meaning of NTS, she stated '*I know about situational awareness but it's not been broken down to the theory* (I1)*',* with comparison and reflection on her own experiences she stated '*I feel not as good as I thought I was, I don't prepare tools or anything I just get asked* (I5)*'* This was not considered in the experiment design and discussed in chapter 12. However, this was one of several acknowledgements and processing of information for the 4 major categories of NTS skills. Understanding between novice and expert were reflected by the participant, with verbalizations of examples to make sense of theory. For example, for shared mental models she explained, '*a good team would have the central part expanded so if I was a scrub nurse I would know what a surgeon's role is, what an anaesthetist's role is* (I6)'. There were 12 cognitive verbalizations overall with 6 in the teamwork and communication codings. Lastly however, there was one sign of the limitation of the material, similar to participant 8 she suggested that understanding situational awareness theory '…*helps me understand but I'm not sure what I'm supposed to do in real-life* (I1)*'*

Participant 14- The participant systematically paused the material when she wanted to talk aloud of her cognitions and firstly reflected on the levels of situational awareness. She stated she had learnt about a similar premise but was within the label of *personal skills*. She proceeded to acknowledge and briefly explain attentional bias with an example. I.e. '*Or some procedure you think is going well but you've forgot to account for something as you were too busy* with *getting it done* (I7*)'.*

Out of the 7 control condition participants, 4 provided one or more statements that suggest they believed there was a limitation to the control condition material. This limitation was surrounding the practice of NTS techniques to improve real-world abilities. Participant 14 also stated, *'I've done it where I meant something else and they meant something else, I think I'd need practice more because knowing is good but it's hard to actually do when you're doing a million things at once in surgery (I13)'.*

Overall, participant 14 used a lot of examples to interpret the information and identified the decision-making content to be the most useful. In the decision-making scenario she had stronger analytical skills than many participants but express being unsure of choices. The decision-making content provided may have helped her to be more confident or understand why her choices were similar to a CAPs processes, therefore were found to be useful to her.

11.9.1 Summary of Findings for Control Condition- During Intervention

The control condition provided insight into how individuals make sense of non-technical training material. Their results were the benchmark for utilisation of information presented. Any alternative NTS training would have to improve upon the positive outcomes verbalized by participants to suggest better potential mechanisms to enhance their skills. Indeed, the main beneficial outcomes for the control condition material were that participants (i.e. 9, 10, 11, 12) acknowledged understanding of most or all of the content, and used examples to apply the data into context they better interpret. The reason for doing this may be to create a mental scenario that included their interpretation of the NTS theory and can be easier stored and extracted in long term memory. This premise was expanded upon in the discussion.

11.9.2 During Intervention- Cognitive Occurrences for Experimental Condition

Participant 1- Out of 8 events in situational awareness, 5 were in the '*process deviation'* category. There were 3 occurrences when the participant also showed their cognitive process in attempting to gather new information from the surroundings, which did not occur in the pre-condition measure, for example stating '*oh, what's that? what is she doing*?'(l2). When paired with the frequent identification and description of the deviations of processes in the scenario this suggested participant 1 was still orientated towards specific cues relating to equipment usage and rule-based practice rather that the social non-technical skills categories. However, the key change was the information gathering at the beginning of the scenario. Feedback for the majority of cues were then presented which lasted 3 minutes and the participant remained silent in this process.

Participant 2- The majority of verbalization of the participant's thought processes were coded into the *comprehension* category of Situational Awareness. Specifically, 6 occurrences were identified to be the participant's understanding of *dynamic processes* of elements within the environment, and 4 were of when these expected processed had deviated (*process deviation)*. For example, the consequential effects of one factor upon another was shown from identifying that '*the surgeon is waiting for equipment which means the scrub nurse is having to be rushed'.* From this, an equal distribution of occurrences was coded across most of the categories of non-technical skills. This may have been an adaptive compensatory behaviour for the attentional bias previously suggested however there was no instructions or advice that stated they had to change their behaviour during the scenario. Overall, during the experimental condition participant 2 had a broader ability to perceive, comprehend, then analyse events with balanced reference to the range of both the cognitive and social NTS elements.

Participant 3- The distribution of coded occurrences of cognition-based non-technical skills classification was also distributed across many of the categories. Specifically, participant 3 had at least 1 occurrence of each category in *comprehension*, and 2 occurrences for each category in *decision-making.* The social categories within the non-technical skills taxonomy division also had a spread of occurrences. This suggested that the participant had a broad perspective of the environment in that they deployed their attention over a broad range of NTS cues and used a combination of strategies to comprehend and predict the scenario. Participant 3 had frequent coding occurrences which detailed the subjective awareness of the cognitive states of others. Her information gathering towards decision-making was influenced by the assumed experiences of the team. For example, she stated *'I think she's feeling a bit stressed because she's got to deal with all this and these girls are just chilling*'(I6), and …*'or someone is just not paying attention 'cos they're just not with it that day.'(I10).* There was an empathic element to the participant's comprehension of the team's behaviour as seen in the scenario. This influenced her decision-making for example she stated '*I would feel bad at saying to everyone 'come on guys concentrate so we can get done'*(I14).

Additionally, when elaborating on leadership of the surgeon the participant continued to postulate the states of colleagues as a main factor in their decision. She verbalized that '*if you're trying to be a leader then I'd want to stay in control and say something, but again is there even any point as they can't help mistakes if they are tired. This would be too hard for me to decide*'(I23). Being described as a higher-level skill, this reflective ability of how individual cognition and action can influence their environment and vice versa, was frequently displayed. The participant acknowledged '*…and I'm still learning the rules of how to be with others when you're just tired (I11).* Overall, during this experimental intervention period the participant displayed a range of cognitive abilities with a team-based approach, with less perception to the processes that related to the prevention of minor undesired events but more of the team's status.

Participant 4- Similar to the pre-condition measure, there was no higher level predictive situational awareness codings and no occurrences of identifying deviations in expected processes. There could be the assumption the participant had these abilities but did not verbalize them, however she explicitly stated she did not know how to predict undesired events. I.e. she stated, *'I don't see anything really I'm not sure how to spot something going to happen* (I4)'. The participant overall used NTS abilities mostly comprising of information gathering and basic perception/comprehension. As with all experimental condition participants the feedback was shown and the participant was asked to concentrate without talking to take in as much of the feedback as possible. It was not as evident what cognitive events were occurring compared to other participant's, however the measure after intervention did capture more verbalizations.

Participant 5- During the intervention material participant 5 showed more than 4x the frequencies for teamwork and communication and situational awareness. The cognitive verbalization transcript increased compared to the previous measures but only had 11 lines of text. The codings were chunked into 2 general sections. This meant that the first 5 lines were primarily perceptual scanning and identification of events, with a shift to comprehension themed cognitions for the remining 6 lines. There was a focus of attention to the tracking of instruments and can be highlighted by the 6 uses of the word 'swab', 3 uses of the word 'instrument', and 7 uses of the word 'count'. These were focussed around the scrub nurse, with mention of this individual 6 times. Therefore, the interaction with the experimental intervention material provided the participant with an unknown factor that allowed her to increase verbalization when compared to the measure before intervention- this was in a real-time manner to what she saw. There was no level 3 prediction codings however, and no mention of leadership actions of the surgeon which matched the pre-condition measure. However, from an equipment tracking process there was a large number of cues identified that were also in the feedback material. There was also note that during the feedback of the scenario the participant was looking around and away from the main sources of information in the video and may have been distracted.

Participant 6- From a chronological standpoint, out of the participant's 56 lines of transcription, the first 10 lines had many information gathering occurrences which suggested the initial scan of the environment. This was also observed in the raw video footage as the participant deployed a fast 'look around' in the VR environment. A few initial data gathering verbalizations were '*What's he doing? The wires are nice and straight'(l2), 'What's she writing on the board?'(l2), 'She's asked for something. She's looking and moving things.'(l3), 'she's just put that on the tray'(l6),* and *'What's she just chucked down on the floor?'.*

Then, a variety of non-technical related occurrences were identified that demonstrated comprehension of an array of patient-related cues. Importantly, the participant not only could identify a cue they believed was non-technically relevant but expanded upon this cue to state their cognitions behind the selection. Indeed, patient safety was also evident to be a concurring underpinning driver towards *how* she verbalized her cognitions.

Diverting away from the dominant focus of *reducing/increasing SA* observed in the pre-condition measure, participant 6 described a plethora of equipment-based events that showed higher level analytical strategy to assess the related cues. For example, she stated '*they have trays which separate all of the swabs... And with her just chucking a blood-soaked one erm there could have been two in that one swab but because its soaked in blood erm you don't really have much depth of perception'(l22).* These descriptions suggested the deliberation of cues with initial rule-based reasoning, but a then more detailed and individualistic justification of how each cue could affect patient-safety. The feedback of cues was provided and the participant could see any differences in cue meanings.

Participant 7- During intervention there were more occurrences of each *perception* encodings, being *of human action*, and *of equipment processes*. Indeed, there were 7 encodings of perception of actions of colleagues, and 6 occurrences of perception of equipment processes. This meant that, like in the previous task participant 7's verbalizations represented their cognitive scanning of the environment to quickly understand what events are occurring. For example, she stated *'the runner's been writing something on the board and the scrub nurse has asked her for something else'(l5),* and '*that scrub nurse has just erm emptied something onto a tray' (l7).* However, the participant did deliberate more on these preliminary cues, and as identified in their decision-making scenario, they had the ability to be analytical of each cue. There were 7 *prediction* based *theoretical deliberations* of events that were patient safety focused. For example, she described that '*the nurse the runners just emptied something onto the scrub nurse's tray without her looking so she might not know where it is erm it can be a problem for counting at the end*'(l8). For most perceptions, there was some form of elaboration which described what typical practice should be, or the consequence of the deviated actions.

11.9.3 Summary of Findings – During intervention

There were more cognitive verbalizations for the experimental condition compared to the control condition and resulted in more NTS related codings. This allowed increased capture of the NTS used whilst interpreting the material presented, especially for participant 5. An unknown factor within the setup of the experimental material provided participants with increased ability to verbalize their thoughts. Possibly, cognitive workload may not have affected them as much compared to the measure before intervention, and compared to the control condition measure however further support for this was required.

Secondly, there was a phenomena of acute cue scanning within the VR environment for 4 of the 7 participants **(**participants 1, 5, 6, & 7). This was different to the measure before intervention and may have been attributed to the VR setup that distributed the environment 360-degrees around them. More effortful deployment of attention to initially acknowledge the layout of the environment occurred which caused their early cognitive verbalizations to shift towards the basic level 1 situational awareness codes in the coding scheme. The effects of this upon improved error related cue identification in VR have not been investigated in the literature.

As previously suggested, some control condition participants used examples to provide a meaningful representation of theory. Further evident was captured to support that without this application of real-life representation, understanding of the concept may not be optimal. Participants 8, 13, and 14 expressed how there was difficulty in converting the information to reuse it and improve their behaviours under real-life conditions. Participant 8 stated, '*I'm not sure how I'd be able to predict things* (I5)'*.* Participant 13 stated*, '…helps me understand but I'm not sure what I'm supposed to do in real-life* (I1)*'.* Lastly, participant 14 stated, '*I think I'd need practice more because knowing is good but it's hard to actually do when you're doing a million things at once in surgery* (I13)'.

This may favour the combined use of the experimental condition material- as individuals receive a simulated real-life scenario that may provide them with more context of how NTS events occur. Within the experimental material, theory was only able to be inferred from the feedback they received towards the end of the session which described the major cues and how a more experienced individual interpreted these. Perhaps a combination of the two conditions may optimally accelerate the interpretation of NTS theory into schema that can be transferred to real Operating Theatres.

However, as this measure was during intervention the summary of findings could only serve as a guide to what may be found when comparing the within and between groups differences in the after-intervention measure. The next section investigated if these differences were present and how Non-Technical Skills towards patient safety were or were not modified due to condition over time.

11.10 After Intervention Measure- Cognitive Occurrences

As the control condition was the representative basis/benchmark of NTS training, the control condition participants (8-14) were again presented first and summarized to generate the outcome of effects. This would better highlight the dissimilarities when the results from the experimental condition participants (1-7) were presented.

Participant 8- Initial verbalizations were towards the reduction of attention of the staff and the communication issue because of this. Certainly, content of most the first 12 lines of the transcript were detailing the consequences of an OT personnel's actions upon their interaction of a tool. This could initially have suggested the participant still orientated their attention to identifying procedural surgical events, the actions of each person, or infection prevention practice but have fewer social considerations. The same method was used in both before- and after- condition measures, in that the participant considered both their knowledge from similar experiences and prediction of future events to explain their reason for cue selection. However, the main change when compared to the measure before intervention was a lot more detail surrounding each cue identified. As with participants' transcripts in the experimental condition, there were more verbalizations in the measure after intervention for participant 8. This was not a profound finding unless changes in the content itself were closer to the cognitive strategies of more experienced individuals. Therefore, upon comparison of the content the measure after intervention held more occurrences of *detailing or justifying a prediction,* and *identifying embedded interactions.*

Detailing or justifying a prediction was in the predictive category and therefore based on the Endsley (2002) model- the highest abilities that can be gained to appraise an environment. After the control condition the participant was able to verbalize more cognitive processes- some of which were at an improved level as they were closer to what type of strategies experts use. These verbalizations captured were categorized by both independent raters as being towards awareness of the environment, as they had absence of any link to making decision. For example, the participant stated, 'i*f she tries to grab something from here, she could easily nick her hand on that'*(l29). However, there was also an increase of decision-related cognitive verbalizations that had the same higher level of *prediction* embedded within it. The decision-based occurrences of *identifying cues* and *describing option consequences* were abilities used to aid towards the optimal choice selection.

An example of this was when the participant described a decision that should have been made regarding tool usage- not only did she verbalize the issue and explain the desired practice, but then elaborates the consequences if the protocols were not followed. This pathway from identification, then decision, then elaboration, can be exemplified in the following broken down exerts from lines 19-26:

'So, I just heard something fall on the floor…' - initial cue perception and identification,

'…and as soon as something falls on the floor, we need to say so…' - recognition-primed and rule-based decision-making,

'*... the runner should come over with some gloves on and pick it up… and then write down somewhere like on the whiteboard…'* - procedure-based actions from the decision made,

'… because that tray needs to be fully complete when it goes back…'- comprehension of consequence in relation to other elements,

'*…they don't know if it is in the patient and will need to look back in the patient potentially.'*- predictive deliberation if incorrect decision/behaviours followed.

Therefore, participant 8 showed an increase in the occurrences of both comprehensive abilities of assessment and improved ability to consider the meanings of cues identified more than coded in the measure before intervention. However, there was still a limited frequency of teamwork, communication, and leadership orientated verbalizations which suggested the control condition may not have aided in the participants ability to perceive a wider variety of non-technical skills, but to better consider the meanings of each cue. This notion was developed further in the discussion.

For the SAGAT ((See Appendix Table 8)) there were 15 correct answers in both conditions- 5 of these answers changed to be less accurate and 5 changed to be more accurate than before intervention. There was no increased frequency for answers that referred to either tools, or to staff therefore these SAGAT changes were not able to be clearly explained as they did not show similarity to their cognitive verbalization results.

Summary of Findings

- There was an increase of decision related cognitive verbalizations, and the participant used their mental models from experience to understand events.
- There was a limited frequency of teamwork, communication, and leadership orientated verbalizations which suggested the control condition may not have aided the participants ability to perceive a wider variety of non-technical skills
- The SAGAT scores had 5 decreased scores but 5 increased scores. The simple count of staff seemed to become less accurate in the post-condition.
- Verbalized more cognitive processes- some of which were at an improved level as they were closer to what type of strategies experts use

Participant 9- Most of the first half of participant 9's transcript was coded in the *comprehension* and *prediction* categories of situational awareness. There was increase in content surrounding the awareness of staff in the video- for example the participant used the word '*aware*' 9 times in her transcript compared to 0 times in the before intervention measure transcript. Although this suggested change in the participant's ability to identify related cues, this caused a reduction in the quality of elaboration. Indeed, as explored in the measure before intervention the participant continued to refer to one cue at a time without elaboration. For example, she stated '*scrub nurse is not really aware of her surroundings, she's aware that the patient is there, surgery is happening, basic things, however she's not aware of the consequences of her leaning on the patient'* (l2). Another example being, '*the HCA still has her back to them, she's not aware, not aware of her surroundings or anything'*.

Therefore, the participant was able to increase the frequency of cue identification but her cognitive explication remained at a basic descriptive level with minimal analytical occurrences. This could simply mean however that the participant did not feel the cues were of more importance towards patient safety and did not want to spend too much time on them. One effect of this increase in cue perceptions paired with minimal descriptive detail was a reduction in *process deviation* coding occurrences. This supported the suggested cognitive verbalization shift as *process deviation* required effortful comparison of violation expectations. Noteworthy, there was an increase in rule-based coding occurrences (from 2 to 5) as the participant verbalized the simpler responses to events which again demonstrated less developed analysis of cue meanings.

There were 10 correct answers for the pre-condition SAGAT (See Appendix Table 9), and 14 correct answers in the post-condition SAGAT. There were 6 answers that increase in accuracy towards the correct answer, and 1 which decreased in accuracy. Five out of the 6 answers that increased were not equipment related, but staff and socially related events. This meant the participant stayed consistent on focusing on the actions of one individual at a time as captured in the measure before intervention. In the SAGAT, one distractor question was '*were the lights adjusted?*' this was added as it was to divert the participants from guessing the specific cues that may be of interest. This cue was assumed to have no meaning, and 12 out of 14 participants did not attend to this cue in both the pre- and post SAGAT which supported this notion. However, participant 9 did identify and provide meaning about this cue, she stated '*the surgeon used the light handle to move the light, he didn't look around before he moved it... He could have hit anyone's head'(l4).* This demonstrated one benefit of being able to identify a large number of cues paired with staff related situational awareness verbalizations as she was able to produce meaning from this cue that other participants did not consider/identify.

Summary of Findings

- A high amount of information gathering and cue identification but with minimal elaboration- this was possibly due to the excessive number of cues/events being verbalized. e.g. there was large quantity, but this compromised explanation.
- The cues verbalized were for larger visual cues in the environment- such as people. This was evident in both pre-post condition measures. It was also reflected in the SAGAT as the questions referring to the count of staff became correct.

- Rule-based factors increased in the post-condition measure, which suggested the control condition material may have primed the participant to develop their ability to comprehend deviations of best practice, by describing the rule break that caused this deviation to occur.

Participant 10- There were 10 less occurrences of coding when compared to the pre-condition measure; 6 less in the *situational awareness* category. This was partially because there was less verbalization, but also because the participant lengthened her verbalizations regarding 2 specific events. These two events were regarding a high workload of one staff member, and the other was the position of a piece of equipment. The participant identified a high cognitive workload of the scrub nurse which involved the initial identification of the tasks being performed ('*erm the scrub nurse looks like she's been asked for something* (l2)', followed by the comprehension that there may be a problem with the scrub nurse *('and she's not quite sure where her instruments are and she's also got a swab in her hand which means she's not, she's trying to multitask* (l5)'. Finally, from this process the participant was able to predict a consequence if the tool was lost. This differed from the pre-condition measure verbalizations which were shorter in length and did not 'walk through' the non-technical event identified as much as the post-condition measure.

With absence from her verbalizations during the control condition, there were also no verbalizations surrounding shared mental models or information sharing to update team understanding. However, similar to participant 8 they may have had difficulty converting this into practical application as it has little visual cues that may have made this concept difficult to apply. For the remaining content in the post-condition measure, the participant also did not verbalize any leadership related cue (3 occurrences before intervention). Overall, the participant showed minor improvement in cue meanings but had shifted identification of events towards the scrub practitioners' actions. Changes in the SAGAT(See Appendix Table 10) questions should reflect this by decreased accuracy for events outside of this field of view.

For any SAGAT question referring to the communication/leadership of the surgeon the participant was correct in both pre- and post- conditions. A reason for the reduction in verbalizations for leadership may have been that she did not consider them as primary to patient safety, but she still had awareness to gain correct SAGAT scores. The correct SAGAT scores before the control condition were 14 correct answers, and in the post-condition 2 answers increased accuracy towards the correct answer. She also attended to more of the closer environment in the video as the verbalizations and SAGAT improvements revolved mostly around the scrub nurse and one trainee who were closest.

Summary of Findings

- The participant verbalized understanding of the importance of NTS and stated the material helped to improve non-technical competences
- Longer details of each cue were verbalized and had mix of perception, comprehension, and prediction. Participant 10 focussed more on the scrub practitioner and the events surrounding that area. This resulted in increased accuracy in SAGAT scores.
- Some NTS elements (such as shared mental models) of the control condition material were not mentioned and did not appear in the pre- or post-condition measures.

Participant 11- There were still a similarly high number of *increasing/decreasing* situational awareness verbalizations paired with explanations surrounding communication. This would suggest no changes in NTS interpretations. The participant identified and repeated the same cue and explanation for the communication event between the surgeon and the runner. However, there were also 3 more occurrences (from 1 to 4) of *process deviation* descriptions that shifted away from communication and instead detailed other cues. For example, participant 11 stated '*it can be quite difficult so they should be paying attention to the vital signs of the patient* (l4)'. Additionally, equipment usage was mentioned which further suggested the participant was attempting to identify a wider variety of events. She stated '*Situational awareness is a big key again and having experience in theatre, swab counts is one of the most important checks so she should have been more aware before just placing it onto the drapes* (l15)*'.*

Therefore, the results had a mix of cue types identified, but still limited within the *comprehension category*. The participant did not show processes of cue combination to form a prediction, or describe how the development of undesired events could occur. There was still a restriction on the type of cues identified and the coded non-technical categories that were used to comprehend each cue. However, the situational awareness theory material in the control condition may have helped the participant to verbalize more of their understandings (+43%), to recognize a small increase in cues and the categories they fall into. Ultimately, NTS were not modified to be similar to experienced individuals but stayed within similar levels and strategies throughout.

Summary of Findings

- Participant 11 largely identified the same cues and verbalized their thoughts with the same perspective and rationale in both the pre- and post- conditions
- There were more occurrences of process deviation, suggesting the participant placed more emphasis on these actions. Expansion to include their possible consequences was not present however.
- The SAGAT scores had increased accuracy for team related questions- the participant had accurate answers. However, with more in-depth elaborations of cues, there was a restriction to the way the participant could utilize the cues towards prediction.

Participant 12- There was a similarity between the pre- and post- conditions as they had matching encodings for all 4 NTS categories. There was less coding frequency during intervention than most of the experimental condition participants and was also very close to being the same amounts as the pre- and post- measures. This suggested participant 12 had a consistent number of verbalizations for each category with no notable changes due to the intervention material. However, transcript content was explored to assert if cognitive process had been modified.

Participant 12 identified the individuals in the video that all other participants also suggested they were '*too close to the sterile field* (I1)*'.* Another similarity was the explicit verbalization of an increased effort to scan the environment. She stated, '*I'm trying to look more but I don't know what I'm looking for really, I suppose it's when it happens, I will know* (I3)*'.* This phrase suggested the material provided knowledge of the concepts, but not context. This concurred with the previous summary of cognitions during intervention. This epitomised the limitation of many current training methods as the absence of application results in such cognitions that suggest individuals cannot transfer/ translate the information to real life.

However, the participant did identify level 3 SA that they did not in the pre-condition measure. Note the sentence, *'the scrub nurse, I've noticed, does hand the surgeon things before he's asked, he just looked up then and she gave him something but he didn't say anything so that would be prediction* (I7)*'.* Therefore, contrary to her concern of what cues to identify, she did show a deeper understanding of the scrub practitioners' actions before the actions of the surgeon take place. The category of prediction increased from 0 to 2 occurrences for participant 12 as she also identified that *'The runners just put something on the tray without the scrub nurse noticing. But she will see something new when she looks at the tray probably* (I11).*'* Between these measures there were little changes in encoding frequencies, yet some underpinning cognitive improvements occurred.

The SAGAT scores (See Appendix Table 12) showed that there were 17 out of 28 (61%) correct answers before the intervention material. After intervention there were 14 correct answers (50%), with 5 becoming less accurate, but 3 different questions being more accurate to the correct answer. Looking at these questions that changed, questions 1, 2, 4, and 6 became less accurate and related to the count of how many people were in the OT shortly before the scenario stopped. One explanation may be that she was deploying her attention to other areas and types of information within the last few seconds therefore did not notice these cues. Therefore, there was a negative change in SAGAT scores, specifically towards the team count on the scenario, with no changes in other equipment related question scores.

Summary of Findings

- Participant 12 verbalized the concepts of the control condition material well. Suggesting she interpreted them accurately.
- She identified cues that were commonly identified both between conditions but also within pre- to post measures
- She explicitly verbalised she was unsure of what cues to look for, however she identified 2 higher level SA cues.

- The post-condition measure SAGAT scores were less accurate for the awareness of colleagues, and may have been due to a shift away from these cues. The post-measure verbalizations concurred this as there was more focus on the surgeon and scrub practitioner and less on the overall team.

Participant 13- There were 50% (6) more encodings which fell into the situational awareness category compared to the measure before intervention. No other verbalizations showed presence of the other NTS usage therefore there were no zero coding. This mirrored the measure before intervention as there were also only 2 encodings. Out of the 12 situational awareness occurrences, the most frequent were from basic perception of equipment (3 occurrences), and higher-level transformation/reflection of cue meaning (4 occurrences). One example of reflection, being closer to introspection, was '*I'm more aware of what people could be doing even if its small, and also how much noise is making issue* (l8).' Overall, the participant identified mostly issues surrounding the scrub nurse along with the positives of communication. This did not change as cues identified surrounded the scrub nurse but there was more reflection on these events, the participant stated '*she's unpacking more swabs because she knows she's going to need them later on in surgery. She's doing a lot I didn't realize how much she's doing compared to the others* (l5). Therefore, the frequency of cognitive verbalizations increased (if only for one NTS category) along with more proactive awareness of what further meanings could be gathered from the noted events.

Lastly, there were 4 SAGAT scores which decreased in accuracy which involved actions of staff. There were 5 increases in accuracy- 3, of which were regarding different staff actions and 1 was the recognition of music. With support from the findings from the verbalizations this suggested the participant's attention did not increase to be more vigilant, but there was more attempt to link cues which is a characteristic of more developed NTS strategy.

Participant 14- During the intervention material, the participant did not increase the number of cognitive verbalizations compared to the before- and after condition measures. This was dissimilar to the experimental condition participants who showed higher numbers during intervention. This meant there was consistency in coding verbalizations for each NTS category even after the intervention material was delivered.

Similar to participant 12 there was suggestion of more attention towards the scrub practitioner and surgeon, which was more explicit for participant 14. There were 11 out of 19 (57%) of the lines in the transcript were regarding the scrub practitioner's actions in some form. Both the coding verbalizations and encoding categories did not notably change between the before- and after condition measures for participant 14. Yet the quantity of verbalizations increased and through this more elaboration of thoughts behind the identification of each cue. This suggested the control condition material may have promoted cognitive verbalization, or promoted more thoughts about the actions in the video being viewed.

An increase in smaller sized cues occurred in the post-condition measure compared to the pre-condition measure. This meant the participant had more focus on equipment and cues that were visually harder to identify due to their size. For example, she stated 'a*nd she also looks like she is looking for a piece of equipment but can't find* (l2)', and 'I *think it might be a swab in her hand* (l10)'. These smaller cue identifications were not as frequent in the pre-measure condition. Although this could suggest the participant attended to cues that had more relevance to error prevention, literature (Mushtaq et al, 2018) suggested communication problems are the most common cause of error not equipment. However, there were verbalizations of how these equipment issues may have been prevented by improved teamwork and communication which was similar to the before condition measure.

Overall, the observed changes in attention deployment and explanations suggested the participant improved their vigilance and NTS towards prevention of undesired events- this was not reflected by the number of encodings. Lastly, the SAGAT scores (See Appendix Table 14) showed that 14 questions were correct in the pre-measure condition, and 16 were correct in the post-condition measure. Additionally, 4 questions had increased accuracy in the post-condition measure (of which 2 became correct) with only 1 question decreasing in accuracy. The 4 improved questions all surrounded the actions or count of staff in the OT. This proposed the participant became more vigilant to her colleagues' action compared to before intervention.

Summary of Findings

- Changes in quantitative encodings from the cognitive verbalizations were minimal and did not reflect the cognitive changes explored.
- Decision-making was more elaborate but the participant struggled to decompress her cognitions to explain the reasoning.
- Before intervention, the participant identified a variety of cues, but did not piece them together to understand how they can create error. After intervention, there were verbalizations of how equipment related issues may have been prevented by improved teamwork and communication.
- SAGAT scores increased in relation to OT staff actions.

Overall, there were improvements in both smaller equipment-based cues but also general OT staff events which suggested the control condition material was the most effective for participant 12 out of the other participant in this condition.

11.11 Experimental Condition Measure after Intervention

Participant 1- There was more explicit acknowledgement of attempting to gather new information from the surroundings as the participant stated, '*So I'm looking around more initially as I want to take everything in first'* (I1). This suggested that the experimental condition setup changed her understanding of how initial information gathering may benefit NTS comprehension for the rest of the scenario. Additionally, compared to the pre-condition measure, there were 4 more occurrences which suggested the participant was using higher-level situational awareness abilities. She was able to deliberate on a piece of information for its consequence in the future and reflect on her own limitations in understanding the events. Indeed, there were no accounts of '*process deviation'* as the participant did not state what *should* or *should not* be performed, but rather the possible reasons for the events to have occurred. This showed the participant tried to extract more details from cues than the initial performance. For example, she commented on a staff member in the video as '*she might be new to the procedure'*, and realizing the potential reason for two anaesthetists when there is typically one- '*actually if there are two, are they swapping over… they are anaesthetists but maybe one is a trainee*' (I10,11). This deliberation suggested that the participant was not only identifying events in the environment as performed before intervention, but applying more of an analytical mindset of non-technical skills to comprehend *why* these cues and events were occurring rather than *what* was occurring.

Six of the 9 SAGAT questions which increases in accuracy were equipment related. As previously specified, participant 1 verbalized instrument and procedural processes more than the social categories in the coding scheme and the SAGAT results show concurrency with this deployment of attention. Correct equipment related questions 6, 9, 11, 12, 17, and 20 suggested that the participant's attention to these elements were increased compared to before intervention. There were 11/23 correct answers prior to intervention, and 14/23 correct answers after intervention.

Summary of Findings

- There was evidence to increase information gathering at the beginning of the scenario. This can improve decision-making and mental model accuracy of the environment for the remainder of the scenario.
- There was an increase in reflection and prediction, however this was limited to being within the same non-technical skills categories. I.e. the verbalizations did not extend to leadership or teamwork, but orientated around the awareness of equipment and individual cues.
- The participant showed transformation of a cue meaning, which is a higher-level cognitive ability and suggested closer cognitive ability to that of more experienced individuals.

Participant 2- The large increase in coding for underpinning cognitive processes was mostly due to an increase in quantity compared to the pre-condition measure data. This may suggest the experimental condition was able to help the participant promote more verbalization, but this would need to also be evident in other experimental condition participants to consider further meaning behind it. Although some socially orientated NTS categories within the coding table increased by 1 or 2 occurrences, it was the cognition-based category of situational awareness which increased the most. This may imply the experimental condition prompted more verbalizations-which was suggested by the coding of 14 more NTS related cognitions. There was evidence the material also helped to make explicit their thoughts better. There were more occasions of '*process deviation'* in the post-condition measure (from 0 to 4) which can aid in error prevention in real life as unexpected cues could be more frequently identified. Moreover, compared to 1 occurrence in the pre-condition score participant 2 verbalized 5 occasions of '*transformation*'. This meant that there were 4 more perceptual and cognitive occurrences of reflecting and reusing information to produce new meaning. For example, she stated, '*has the music gone up? Maybe it's just me paying more attention to things but it sounds like its more distracting'*(l16). This is a cyclic process of factoring in the effects of the participant's own perceptions to find how cues can be interpreted in a different or more accurate way. An example of promoting explication of cognition was '*so, from what I know now it's not so much what they are doing but also how and why they are doing these things'(l1).* This quote suggested there was similar relevant events identified in both conditions, but in the post-condition she further attempted to expand on *why* they believed they were important and *how* they related to error prevention etc.

Although there were suggestions of improved NTS, there were minimal improvements of accuracy for the SAGAT answers (See Appendix Table 2). Participant 2 correctly answered 16 of the 23 questions after excluding 4 distractor questions, in both conditions. This meant that although the participant was better able to deliberate on the meanings of the cues identified, their ability to accurately identify information and cues was not positively affected.

Summary of Findings

- The participant had an analytical decision-making ability before intervention, and most SAGAT scores were correct in the pre-condition.
- There was more identification and elaboration of *process deviations* in the post-condition measure-suggesting improved ability to verbalize and describe *how* these events can affect the OT.
- There was a reduction in *perception* related description and an increase in *comprehensive* explanation which concurred with the previous suggestion that improved ability to convey their cognitions of the events had occurred.
- There was more awareness of particular problematic cues which could help to address minor issue if transferred to real life events.

Participant 3- Similar to the pre-measure coding results, the participant was able to identify events relevant to the operating theatre. However, there were less basic perceptual verbalizations of behaviours (from 11 to 5) and instead an increase in gathering and piecing together information. Indeed, comparative to the previous statement in the pre-measure condition of stating 'I *don't really think I've seen a risk of an error yet'(l8),* the participant identified a key equipment tracking event that characterized the procedural thought processes to avoid error. Firstly, there was the basic perception of the staff members actions - '*the runner has been watching what's been happening, so she actually has seen the 2 swabs in the bag'(I13).* Then upon comprehending the meaning of this, she applied the implications towards her idea of the desired behaviour - '*and they need it crossing off the board'(I13)*. With final recognition of the main patient safety orientated goal- '*to keep the count accurate'(I13)*.

This change may be encouraged due to the experimental condition having some form of priming effect. Similar to participant 1 who stated '*So I'm looking around more initially as I want to take everything in first'* (I1), participant 3 also verbalized more attentional deployment towards gathering information. These were '*I'm just looking around in case there is anything else(I15), and 'last time I wasn't looking there'*(I7). This increase in scanning was concurrent with the increase in SAGAT scores (See Appendix Table 3).. Participant 3 correctly answered 10 of the 23 questions in the pre-condition SAGAT, and 18 in the post-condition SAGAT. There were more information gathering encodings and more awareness of the meanings of the cues towards minor problems. Indeed, there was no trend in the type of improved answers as they were instrument, event based, or colleague cues, this may have been partially caused from the participant's increase in information gathering as evidenced in the encoding.

Summary of Findings

- Participant 3 was initially was not able to link and elaborate on events to analyse if problematic events may occur however basic perception frequency decreased, information gathering verbalizations increased.
- Improved elaboration of certain cues via cognitive verbalization, demonstrated more considerations for each cue meaning and its effects.
- Reflection on performance from the pre-condition measure helped performance when in the post-condition measure.
- SAGAT scores generally increased in accuracy and frequency for all types of NTS related information, suggesting a stronger distribution of awareness even though coding frequency decreased in the post-condition measure.

Participant 4- Participant 4 identified the majority of events verbalized in the pre-condition measure. For example, she mentioned the effect of music and was also captured in the SAGAT. However, there was an additional event spotted that had commonly being identified amongst other participants which suggested increased scanning of NTS related events. Importantly, the participant expanded on the initial identification to explain the consequences. I.e., '*It's not affected anything but if maybe the staff were more aware then they could have stopped her just in case* (l8)*'.* This demonstrated the participant could link events to find a preventative measure for a problem. This did not occur in the pre-condition measure and was a suggestion that the feedback from the intervention material aided the participant to think about consequences of cues more in the environment. However, cognitions that refer to leadership were not found in participant 4's verbalizations during and after intervention. This suggested the intervention material did not promote inclusion of identifying leadership cues in relation to error avoidance. Additionally, the SAGAT scores (See Appendix Table 4). did not reflect this potential increase in cue scanning as the score did not change. There were 12 correct answers in the pre-condition measure and 12 correct answers in the post-condition measure. However, accuracies between questions changed so that 25% (7 question) of the questions increased in accuracy, but 25% of different questions decrease in accuracy.

Summary of Findings

- An increase in the suggestion of consequences of certain cues, suggesting further elaboration
- No encodings with the NTS of leadership embedded within them
- Large variations in accuracy increases and decreases, this suggestion a shift in attention which improved some scores but meant attention was not deployed as well to other cues
- A larger spread of NTS encodings (except leadership) that suggested the material stimulated performance to understand the environment.

Participant 5- The cognitive verbalizations reduced when performing the after-intervention measure compared to during the intervention itself. There were 8 lines of encoding which included 3 new NTS categories of *transformation/reflection.* There was also a shift in the context from equipment related cues to individual related events in the post-condition measure compared to the pre-condition measure. For example, she stated '*Which I didn't realize there are like 3 groups not one big one because I'm not really there, I kind of see it more how the people are divided up and I'm just stood alone but watching the room* (l7)'. Overall, with the minimal cognitive verbalizations in the decision making and pre-condition measure there was limit to understanding NTS ability before intervention. Although cognitive verbalizations, therefore capture of more NTS categories, were evident during intervention this was not maintained in the post-condition measure. As a consequence, the comparative information that can be asserted was that the participant shifted focus from equipment related cues to team and socially-orientated cues after they had received the intervention material.

For the SAGAT scores (See Appendix Table 5), there were 11 answers that increased in accuracy in the post-condition compared to the pre-condition measure (with 7 becoming correct). This improvement was more than 80% of all participants. These were a mix of staff related, equipment, and awareness related cues. However, 3 were less accurate and a 'swap' in accuracy occurred to some magnitude. Therefore, there were 8 correct answers in the pre-condition measure. This meant 15 correct answers in the post-condition measure from 11 correct before intervention. Notably, several were regarding the staff processes and it was highlighted the participant shifted attention in the post-condition measure to such cues.

Summary of Findings

- There were minimal cognitive verbalizations in the pre- and post-condition measures yet during the intervention material the verbalizations increase, and occurred within several other participants.

- There was absence of higher-level situational awareness cognitions but suggestion of higher-level decision making. The minimal verbalizations hindered the exploration of these factors.

- There may have been an unknown variable that prevented the participant from verbalizing their cognitions. Of those captures, a shift towards social NTS were predominately present.

- The SAGAT scores accuracy increased by more than 90% and was the most accurate of any participant. This suggested increase SA towards the cues mentioned in the SAGAT even though cognitive verbalizations mentioned only few of the same cues.

Participant 6- During the experimental condition, the cognitive verbalizations within the Verbal Protocol Report produced 63 occurrences of underpinning cognitive processes. This was the most across all participants and all parts of the experimental sessions and was approximately 4x the number of occurrences compared to the pre-condition measure. However, there was not an increase in 3 of the 4 categories in the post-condition measure which stayed approximately the same. This suggested high frequencies of verbalizations during intervention were not maintained in the post-condition measure.

Although the frequency of categorized verbalizations decreased after the experimental intervention, higher level non-technical skills were observed compared to the pre-condition measure results. The participant displayed a new cognitive strategy not coded in the previous measures. This was the awareness of the limitations in their comprehension as there were 4 coding occurrences of *transformation/reflection* which denoted the participant's meta-awareness of their previous meanings of a cue. This suggested that the experimental intervention provided the cognitive scaffolding requirements for the participant to have this development in cognitive ability. Further evidence was the identification of a new cue was discovered by the participant that was not spotted by others. This required more developed cognitive and social skills as this described event has no physical cues in the scenario and was purely conceptual. The participant had to simulate the surgeon's perspective to then assess if they were aware of the anaesthetist. I.e. she stated, '*I don't think I noticed in the last video, but the does the surgeon even know if the anaesthetists are there or not. I'll keep an eye on and see…. but looking now, he could do with just looking around'(l4).*

Another example of '*transformation/reflection'* was that the same cues identified in the pre-condition measure were transformed to show new value. Instead of being related to staff purposefully reducing their situational awareness, (e.g. *she's positioned herself in a way that she can't see what's going on around her (*pre-condition measure transcript l9), the participant analyses a similar cue in a bigger context and transforms the cue meaning to be surrounding fatigue/tiredness (e.g. '*It must be over soon because people do look fed up and stretching a little, she has her back to the surgeon as she's resting and I didn't notice before, a few stretches from the observers there').* This suggested the experimental intervention may accommodate for the promotion of further deliberation of certain cues, when compared to the pre-condition measure.

The experimental condition did not explicitly include theoretical explanations of non-technical event or skills but instead were inferred from the scenario events and feedback themselves. Participant 6 acknowledged that this format allowed patient safety-related practice, and a certain amount of inference to the theoretical elements linked to each cue/event. She stated there were positives to '…*follow each person of what they are doing and their effects, and I know it links to how we think and work together and think as a whole'*(l9).

Before the intervention material, the coding results mostly captured the participant's ability to understand the changes of situational awareness of individuals in the scenario. She also frequently identified the deviations from processes they were expecting to occur. However, after intervention there was a more distributed capture of coding results (similar to participant 4). This had parallel with the SAGAT scores (See Appendix Table 7), as there were 18 correct answers in the pre-condition measure, but 6 answers that were more accurate in the post-condition measure (excluding distractor questions). The 5 improvements in accuracy were between both equipment and staff events which showed general improvement, in more than one type of cue.

Summary of Findings

- Participant 6 reflected on her previous perceptions to reinterpret cues, which displayed self-awareness of how cues can be interpreted differently at different times.
- Many identifications of reducing situational awareness occurred in the pre-condition, this was reduced, and a spread of cognitive verbalizations was captured in the post-condition measure
- There were 5 more accurate questions in the post- condition SAGAT but 3 less accurate questions. This was parallel with most participants regardless of condition as between 2- 4 more questions improved in accuracy than questions that reduced accuracy.

Participant 7- There was reduction in encodings for three of four categories however verbalizations for teamwork and communication increased by 5 occurrences. Three of these were surrounding the participant's verbalizations deliberating upon other's cognitive perspectives. For example, she stated '*I think she recognises different stuff from a surgeon's perspective than someone that scrubs and runners'* (l15). This category was not coded in the pre-condition measure or during experimental intervention and suggested an attempt to gain a shared understanding that the staff in the video material may have. Furthermore, there were 9 correct SAGAT scores in the pre-condition, and 14 correct SAGAT scores in the post- condition (See Appendix Table 7). These 5 improved cues accuracies were staff related, more than equipment related. For example, questions 3, 5, 13, 17 were the activities of individuals. This concurred with the participant's increase in teamwork and communication coded verbalizations.

The measure after intervention had more content than the measure before. This was not predicted but meant that participant 7 provided more *comprehension*-based verbalizations (from 2 to 6 occurrences), which meant that their verbalizations changed from perceptual identification to more immediate explanation of cues in relation to patient safety. This would be classed as level 2 situational awareness. For example, the first line of the transcript read *'the blade is being passed without being guarded, but for her that might be just the way that she likes it she might not be bothered about the risk of harm' (l1).* This suggested that although the theoretical elements of this non-technical event were not presented during intervention, it may have aided the ability to develop attention to cues in the environment without understanding their theoretical underpinnings.

Summary of Findings

- Participant 7 improved attempts to ascertain the perspective of staff to understand their actions better
- There was a shift towards teamwork and communication, paired with more comprehension of these action in the scenario
- There were some increases in SAGAT accuracy but also equipment related decreases, this would support the shift to socially- orientated events.
- The participant had analytical decision-making strategy- this was only evident during the experimental material presentation but not in the pre-post measures. Verbalizations were orientated to the actions and consequence more than reasoning behind each action.

11. 2 Interpretative Phenomenological Analysis of Reflections

NVivo's (NVivo, 2020) explore functions allowed in-depth assembling of final themes (See Figure 40). The first of two overarching themes was called *Awareness of own limitations*. It divided into two 2nd level themes. The first theme was Reflection of Limitations, which had 1 bottom level theme within it, called Specifying Limitations. The second 2nd level theme was Limited Awareness of Event Link(s) and did not have a further level.

The second overarching theme of *Effects of Intervention* divided into 3 2nd level themes, being Cognitive Overload from Intervention, Improvements in NTS, and Negative Effects of Usage. Both Overload from Intervention and Negative Effects of Usage did not have further theme levels. However, Improvements in NTS has 2 bottom level themes named Improvements in Comprehension/ Prediction, and Improvements in Attention.

Limited Awareness of Event Links

This theme surrounds the participants' reflection of either their overestimations of abilities before the intervention, or their absence of understanding of full cue meanings. The most representative quote for this theme was from participant 1 who stated, *"I didn't see the, well, I might have seen the actual behaviour but I didn't know it was connected to something else" (I1).*

The pedagogical nature of the intervention meant that feedback was given after participants had attempted tasks for themselves. This provided them with a comparison they can reflect upon better than being shown an example before attempting the tasks. An example of this comparative feedback in participant 2 was, *"I wasn't that far off in terms of the number of things I identified, but there were a few links that I didn't see", s*uggesting the structure of the intervention material helped them to better reflect on their performance.

The benefits of this new perspective to their own cognitions were supported by Korovin et al (2018) who suggested experienced surgeons spend more time engaging in metacognitive activities. Metacognition is a key element to expertise which is not limited to richer mental models, heightened awareness of erroneous events, and more detailed understanding of the dynamics of the OT. This theme was the most impactful towards the aim of enhancing NTS towards increase patient safety as it facilitated metacognitive comparisons for users. Because if this, higher level decision-making may be produced. For example, one participant showed an awareness of their own limitations in responding to a cue. The participant stated, "*she said so much more on the best option, or not more but just better reason that, I felt I knew but I thought 'why didn't I say that if I knew it'-P6*. This is a complex metareflective ability as not only does the participant become aware that their choice of responses may have not been optimal, but questions on the possible reasons why her understanding and subsequent decision-making strategy led her to those actions. As stated, these are key factors to form expertise.

A negative consequence was noted for one participant who did express a reduction in confidence due to missing 1 or 2 key cue meanings, but stated the overall benefit received. Other example of participants beginning to reflect on their limitations were, "*I did learn about non-technical skills even thought I felt I know them*"-p3, and "*you can reflect on what you've seen and how you and you can learn from that and when you're in theatre you can you will it will help you maybe just think about it a bit more*".

Attempts to improve prediction and reinforce safety priority

This IPA exploration complimented the quantitative findings in suggesting that within the one hour session using the intervention material, participants improved their cognitive abilities in NTS. The indications of this are extracts from participants that suggest development of either comprehension, perception, or prediction. These changes are drivers towards increasing the barriers against undesired events, when detailed in the Swiss Cheese Model (see 5.4.1). Participants behaviours attempt to rectify minor issues to prevent the alignment of multiple problems which can combine to impact upon the patient. Cristancho et al (2016) suggested such cognitive processes are effortful 4th level of thinking and concurred with Korovin et al (2018) that this continual process of reflection contributes to the advancement of understanding in the same process as experts utilize.

For example, participant 6 showed effort in predictive ability- "*It was good to be able to practice looking at the room and just pausing to guess what will be next. I was doing it a few times, trying to just say what is going to happen*". This quote suggested the participant changed their epistemological worldview- the intervention material accommodated this practice. Their change in epistemological stance before and after the session was inferred by allowing flexibility in their justifications required for their perceptions to be considered accurate. The sources of knowledge and beliefs can be explored to find if modification of their ontological concepts could improve their worldview and experience of operating theatre events. The relations between the concepts and categories in the scenarios were given different meanings by participants.

Participant 2 stated, "…*but it has shown me how other factors can actually lead up to some impact on what I'm doing, not just small things like waiting for a tool for a few seconds longer, but it helps to show how my ability to keep an eye on the situation, if I can do it or learn to do it, can help me to lead others and keep control of things to steer the direction as we would like*".-p2. The control given to participants allowed them to go at their own pace in the exploration of events and help them to practice predicting events. In doing so, their knowledgebase was questioned and when simultaneously occurring with meta-reflection, participants successfully reflected on their abilities to predict/expect based on the cues they identified- i.e. "*So, it made me think more why I was looking at something and what it means"*-p6.

Specifying and Reflecting on Limitations

As the experimental intervention can direct attention to cues and add information about the cue in terms of its meaning within the system, participants' reflections suggested they had become aware of their limitations in their perceptions/assumptions of a handful of cues. Firstly, direction and feedback are provided, and it was this that allowed reflection on their cognitions. Participant 3 stated, '*I did see some things the second time because the use of this kind of directs me to look at things.'* Participants have reflected on their own cognitions based on the feedback and questioned their assumptions of these cues. Upon reflection some participants identified not only the limitations in these assumptions but pinpointed *why* they had occurred- this was not expected in the research aims. For example, participants 1 stated *'But normally someone points it out, so I didn't really think about it. There was a lot to take in, but normally if someone puts something onto your tray it's because you've asked for it.'*

This suggested that the experimental material provides the participants with metacognitive assessment that may be translated into real world practice. Indeed, real world examples were used to self-reflect. For example, participant 2 stated, '*I think that's probably because I was distracted myself on the main event which was trying to see errors that were happening on the table…I sort of have to very much tune out almost to all the external distractions coming from the runners and the other scrub nurses. So maybe that's a product of maybe why I sort of ignored that part of it.* Participant 2 digressed to summarize that, *'it has shown me how other factors can actually lead up to some impact on what I'm doing'.*

Most participants had some element of reflective comparison of their assumptions of cues/events on the scenarios. This aligned with earlier descriptions by Flavell (1979) that metacognitive skills can be improved from reflection and regulation of memory, performance, and comprehension. "thinking about thinking" (Flin et al., 2008) is typically performed in feedback sessions to promote development of metacognitive abilities. KirkbakkFjær et al. (2016) suggested self- reflective debriefs aid in improvements in non-technical skills in mental health nurses. This finding was not considered in the project as the session was short, and it was predicted participants may not be that efficient at verbalizing what they have seen/thought but shows their effortful interactions in trying to learn when using the intervention material.

Cognitive Overload from Intervention

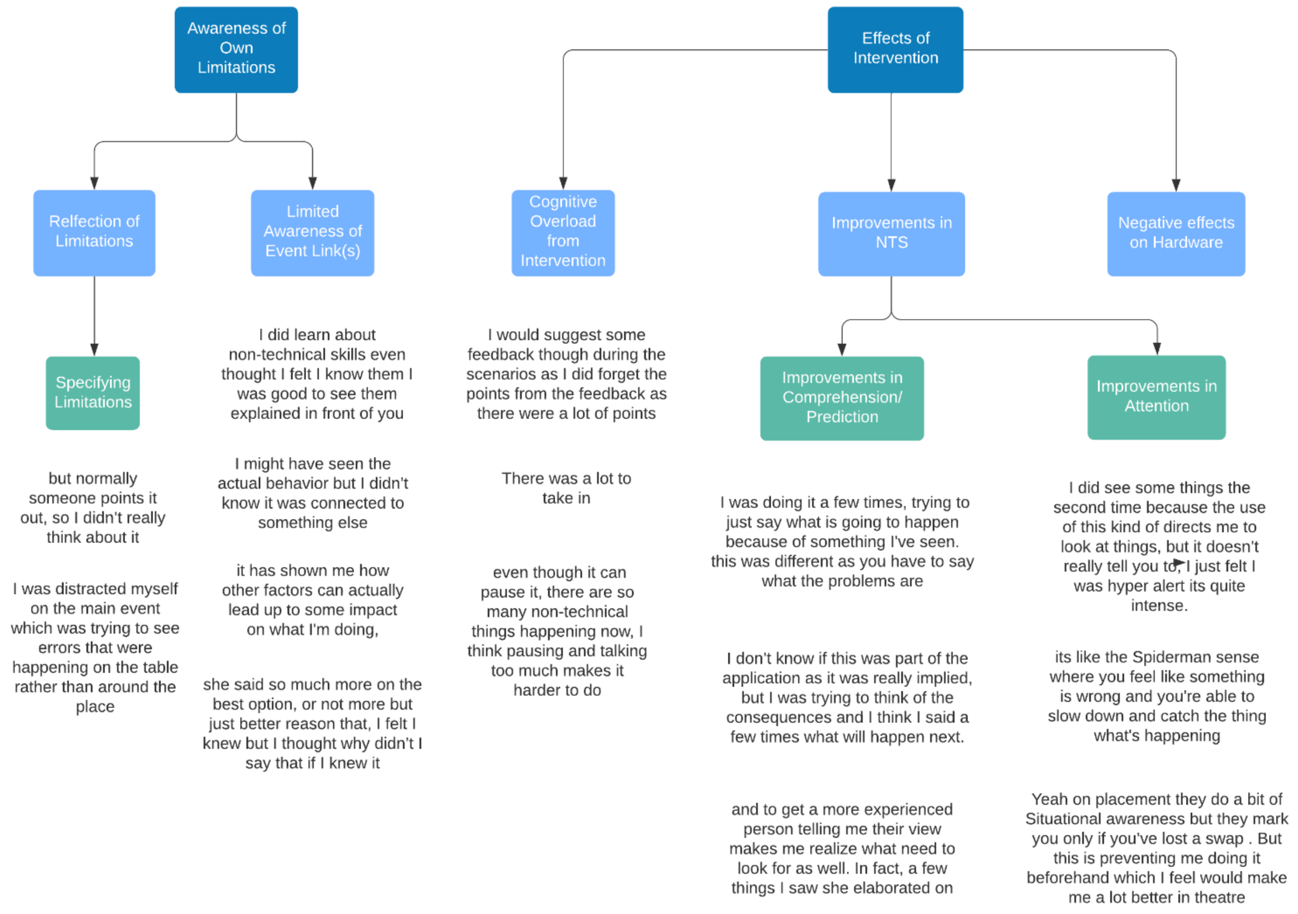
There was a common negative experience distributed across participants. The new material novelty, and cognitive requirements of the intervention material had negative effects on pedagogy at times. Experiences orientated around being too many 'things' in parallel. Their words such as 'a lot', 'so many', and 'too much' highlighted how their experiences had times where their working memory were experienced to be full in that no new information could be taken in. Therefore, there was inference their performance may have been affected. The design of the intervention material allowed pause of the videos; however, this was an unpredicted event when the number of stimuli were too much even in a short video < 2 minutes. The consequence was that some participant forgot what cues were being referred to when they watched the feedback condition. Additionally, although one participant did pause the video, they explained they were motivated to talk, but talking for too long made her forget what other events were occurring. This theme was important in understanding how participants experiences were similar if large bouts of information caused reduction in ability to interact, followed by (if they paused the video) a reduction of their working memory of events if they focused on one event for too long. This wave-like frequency of performance/attention was a phenomenon that detracted away from the participants goals and a key reason why this was a phenomenon they produced through their worldview.

Improvements in Attention

There was a common theme of participants expressing particular cues or actions that they were directed to, or attempted to look for. They described how the intervention material both implicitly and explicitly guided them to items such as equipment, communication etc. Their perspectives were affected by the intervention material in a sense that they experienced their attention being manipulated in order to identify information and resulted in increased vigilance for some participants. One participant described how it made them feel like Spiderman, which was interpreted to suggest their proactive scanning of the environment to detect cues, in a similar premise of the concept of 'Spidey Sense'. This has been synonymised with high levels of intuition and emotional intelligence in Naturalistic Decision-Making research (Eason, 2019). Contradiction of this theme is in the SAGAT scores as the majority of findings do not suggest improved attention to cues, however the consistent references to *how* the tool directs attention has positive outlook for improvement for more use than a single session. One participant stated ‘the use of this…kind of directs me to look at things’ and describes increase in attention in identification of such cues. Overall, a positive description of the participants experiences of the effect on their attention were expressed. The phenomena of boosted attention seemed to be perceived as beneficial by participants, however this does cross over with the previous theme of Cognitive Overload where the areas of stimulation are too great.

Figure 40

Diagram showing the ordinal and sub-ordinal theme found by the experimental participants.



Chapter 12- Discussion

12.0 Aim

The aim of this thesis was to explore the effects of two types of training materials on surgical trainees' non-technical abilities- based on the cognitive and social aspects of criteria initially formed by Yule, Flin, Paterson-Brown, Maran, and Rowley (2006). These criteria were then developed mainly by Flin et al (2006), Brennan (2006), and Yule et al (2008) however the theories and models involved stem from several domains including aviation and military. The aim was to assess the impact of condition on the NTS mechanisms used to perceive, comprehend, and predict scenarios. This was along with changes in ability to analyse meaning of cues with more comparability to experienced practitioners. From this, the evaluation of effect from the two conditions could suggest how NTS are being improved along with the specific components within a condition which may be attributable. To perform this exploration both qualitative and quantitative methodology was used to identify where significant differences occurred between conditions, then investigate underpinning cognitive changes in NTS usage as represented by participants' cognitive verbalizations.

Several rationales formed to generate the opening in research for investigation in this thesis. Statistics from patient safety assessments and error prevention along with arguably outdated training materials were key drivers. Accessibility to resources that were suited to future healthcare professionals was also a driver to create and evaluate a contemporary tool to improve undervalued non-technical skills used every day for all areas of healthcare.

The perspective for investigation into this subject area and goals of the thesis concurred with Anders Ericsson's (2015) review on the acquisition and maintenance of expertise. Ericsson (2015) provided overview of the limitations in traditional teacher-guided practice and compensatory continuing medical education materials when real-world performance was the resultant measure of success. Indeed, he stated upon review that accumulated hours of professional experience from '*attending CME lectures did not affect any meaningful changes in participating doctors' actual practice'* (Anders Ericsson, 2015, p1475). Although this encouraged technical skills simulators to be created that had positive results upon real-world performance, there was significantly less recognition of the need to match NTS acquisition for optimal patient safety. Anders Ericsson (2015) suggested two strategies- *designing learning environments with libraries of cases, and creating opportunities for individualized teacher-guided training*. This should allow motivated individuals to acquire refined mental representations. Few comparative studies exist for the measure of value from virtual reality NTS training in operating theatres to solve the specified issues. Therefore, there has been difficulty and limitation to acknowledge the reliability, capability, and future practicality of this emerging technological methodology. However, this thesis may be a starting point for further development into NTS training with VR simulations and human factors-based techniques for improved expert-novice training.

12.1 Transformation and Reflection of Information

There was a significant increase ($p= .002$) in occurrences for the experimental condition (0 to 2.43) compared to the control condition (0 to 0.14) in the processing of cues in a different way that resulted in new information being gained from the same cue. This was the *transformation/reflection* coding category within the broader NTS social skills range and was used as a code stemming from the 'Reconciliation cycle' by Cristancho et al (2016). Cristancho et al (2016) suggested expertise is a '*continual process of consideration and refinement, and that refinement contributes to the creation of new meaning'* (p6).

On average the experimental condition participants demonstrated a refinement of knowledge to create additional information not previously existing. The underpinning mechanisms were similar to CAP2 as discussed in 6.7.3 of chapter 6. CAP2 demonstrated a high-level cognitive ability to control time and readjust current understanding by transforming previous information, thus re-establishing a more accurate projective (level 3) situational awareness. In parallel, an example from participant 6 showed her initial interpretation of a staff member was negatively viewed, but after intervention interpreted it differently and in a wider context to the rest of the environment. She stated, '*she has her back to the surgeon as she's resting and I didn't notice before, a few stretches from the observers there'*

The experimental condition was better suited to enable participants to intuitively reflect upon their non-technical abilities and reuse cue information to explore different meanings. This has been suggested to be a cognitive ability used by experts (Cristancho et al, 2016). Whereas the control condition material was representative lecture-based material which may not be supportive of the same cognitive processes during the transfer of knowledge. Participants could reflect upon what they understood about a theory during presentation, and how this new information may change their perspective- but there may be limitation to develop different cue meanings to real-life scenarios. Participants would have to reflect on their performance in real-life concurrently to applying new knowledge. This may not be as conducive as the experimental condition setup which can alleviate this workload.

Certainly, there was concurrency with this premise, as the control condition participants specifically attempted this method during their intervention material exposure. With verbalizations from most participants on their understanding of the content being displayed, many control condition participants then used examples or past experiences to apply the new information into context they better interpreted. They were able to take information that was only conceptual and break down each component, convert this to meaningful practical understanding, and verbalize how their NTS may benefit from the material. This was a representative goal of current NTS workshops and lectures as they aim to support individuals to understand the human element within a system, and how their own cognitions and actions can affect a system. Therefore, based on the aim and prediction of this thesis this was the first indication that the experimental material setup had advantages over the control condition. It may aid in the acceleration of trainees' non-technical abilities by providing them with transformative cognitive mechanisms that more experienced individuals draw upon.

For both these conditions the same underlying cognitive process was occurring to make sense of this information. Expertise can be created from the effects of experience which aid the formation of accurate and robust mental models -knows as *categorization mapping* of a wide variety of scenarios (Endsley, Bolte, & Jones, 2003, p24). This can include accurate and robust models of non-technical cognition/behaviours. This may be the same process for participants as they attempted to strengthen their schemata (a representation of a theory;

Lipshitz & Shaul, 1997) which included non-technical elements. The advantages of this occurring in simulation training is observed when in a real scenario as an individual will use pattern matching to find the best fit between the practiced representations compared with the cues in the existing environment they are in. Although not investigated in this thesis, the scripts of an individual would then be retrieved after identification of the best match of scenario available. Scripts have been defined as pre-existing knowledge structures and consequences regarding an event (Eysenck & Keane, 2005, p383; Farag et al, 2010). There was a chance that all participants' script of the scenario after intervention was modified by a factor of the condition they were in. What could be specified was that the experimental condition setup masked the theoretical information presented, which changed the processing required to convert and store the units of knowledge that comprise of a schema.

12.2 Attention to Team-based Cue Types

There was a significant increase in cognitive verbalizations for *shared mental models* for the experimental condition compared to the control condition at the two measures of time (p=0.018). This meant there was a main effect of condition on the participants' attempts to share the same mental models of individuals within the after-intervention scenario. This was indicated by the encoding of cognitive verbalizations which highlighted this NTS ability within them. Secondly, and with overlapping NTS mechanisms the code of *Taking Perspective of Others* suggested a significant increase in cognitive verbalizations for the experimental condition compared to the control condition at the two measures of time (p=0.007).

The implication of these findings was that participants attempted to find the perspective of others by estimating the cognitive representations of the staff on their current tasks. For example, after intervention participant 7 stated, '*I think she recognises different stuff from a surgeon's perspective than someone that scrubs and runners, erm like with the crossing off the documentation at the end*'. The participants had significantly more verbalizations which suggested they had increased their attempts to overlap their understanding of the environment with staff to enhance their understanding, more than the control condition participants. This has been shown to help to describe, explain, and predict the environment (Mathieu et al, 2000). Zheng, Swanstrom, and MacKenzie (2007) investigated the impact on team anticipation, being that colleagues anticipated each other's needs or action and helped to prepare for the events before they were executed. Anticipation for each other's actions displayed more efficient communication and suggested that the participants involved had substantial measure of predictive situational awareness that enabled them to accurately prepare for a task they believed was going to be required.

One explanation for this finding may be attributable to the VR equipment usage. The design of a VR headset is different to 2d television screens or PC monitors as the same image or video is stretched around the user in a 360-degree sphere. The effects are that increased head movement is required to assess the same scenario when compared to a 2d screen. Eye movement when viewing a 2d screen can cover most of the surface with minimal head movement where required. The primary support for this was the cognitive verbalizations from 4 of the 7 participants (participants 1, 5, 6, & 7) during VR usage who expressed an acute cue scanning phase within the environment. More effortful deployment of attention to initially acknowledge the layout of the environment occurred – this caused their early cognitive verbalizations to shift towards the basic level 1 situational awareness codes in the coding scheme (*see section 10.7.1*). Note that although the control condition had different material during intervention, it was the significant finding in NTS codes pre- and post- intervention for the two NTS categories that was displayed, due to the effects of VR during intervention.

The overt actions of effortful cue scanning may have been due to the perspective of first person in VR. As the field of view is stretched around the user, cues become larger when compared to a 2d screen (except for CAVE systems, *see section 1.3).* There may have been an undesired advantage towards certain cues that when apparent in 2d, become significantly more apparent in VR. Socially orientated NTS cues would most likely revolve around people therefore when combined with the first-person perspective these cues become prominent in comparison. The greater shifts from equipment related cues to team related events were displayed for participant 5 and 7. An example which epitomised this notion was from participant 5 who stated, '*…which I didn't realize there are like 3 groups not one big one because I'm not really there, I kind of see it more how the people are divided up and I'm just stood alone but watching the room* (I7)'.

To suggest this could have a negative effect on NTS abilities, therefore error avoidance, would appear antithetical as this VR setup would better represent the perspective in real-world settings. The feedback in the experimental condition may have highlighted the usefulness of this system as participants were 'dumped' into scenarios in their pre- and post- measures. They were not given any information about the time, type, or status of the surgery. This suggested the experimental condition may have promoted participants to adapt this scanning strategy by evaluating the mental processes of the staff as a shortcut to better make sense of the scenario.

However, there has to be caution in diverting attention (accidentally or purposefully) to certain cue types as this implies less attention may be deployed to another type of cues (i.e. equipment, patient stats etc). The immersive effects of VR may draw user's attention without their realisation if they are not informed of this. Ironically with no prediction of this occurring, this thesis previously stated that '*Participants were able to see most of the detail in the environment however specific writing was not legible'* (see section 8.*4)*, and demonstrated caution upon the effects of VR upon cue attendance.

This concurred with similar research which highlighted that use of VR can promote the mechanism of shared perspective and awareness to increase empathy, pro-social behaviour, and induce helping behaviours. However, there is no research on the usage of VR to improve shared mental models in the OT team. The findings that participants attempted to be more aware of the cognitions and behaviours of other staff in the experimental condition may be able to address an issue raised by Gjeraa et al (2019). Gjeraa et al (2019) assessed shared mental models of 172 team members in operating theatres by pre- and post- operative questionnaires (similar to the SAGAT in this thesis). Gjeraa et al (2019) suggested that although concurring with risk assessment procedures, the operating room nurses were the least aware of each other's and the surgeons' problems. Further research into the use of a VR tool based on this study may uncover the extent to which mental models and schema can be aligned in virtual scenarios.

12.3 Meaning and Impact of Findings from IPA Analysis

To explore if participants experienced similar underpinning events IPA was applied to their data. Their data represented experiences that were significant enough for them to verbalise and share as part of their overall experience of the condition, The theme *Limited Awareness of Event Links* was extracted from participants data to highlight their retrospective performance and help to support the design sup pounding deliberate cognition as suggested by the guidelines of deliberate practice suggested by Ericsson and developed by others (Ericsson, Prietula, & Cokely, 2007). Deliberate cognition in the theme limited event awareness is was may be missing from current resources available to students. The first-person perspective and personalised usage facilitated specific mismatch of their own actions. The possible negative outcome of use based on this theme may be generation of over-vigilance in real-life events. The participants evoke their acknowledgements that they may not be fully comprehending all events- and as real-life dictates, they cannot go back to a point in time as compared to VR, they may become hyper-attentive. Further use of the intervention material should use caution as the high-fidelity combined with personalized comparisons could create anxiety or excessive compensatory efforts to improve their performances if future users also have the same insights found for the theme Limited Awareness of Event Links.

Limited Awareness of Event Links theme and the *Specifying and Reflecting on Limitations* theme may be interpreted as overlapping; however, they were distinct themes. Explicit awareness of their limitations suggests a shift in the ontological and epistemological perspectives, and deserves a separate phenomenological theme. Then, *Specifying and Reflecting on Limitations* is founded and developed from limited awareness theme. Although developed upon, it has significantly different processes occurring, being that the participants tried to specify and justify the cognitions they had which limited their understanding of certain cues. This is a difference phenomenological happening for individuals and is a high- functioning ability. The theme of *Specifying and Reflecting on Limitations* cannot occur without presence of *Limited Awareness of Event Links, an*d although separate phenomenologically, they are both needed for the skill acquisition desired. The more occurrence of *Specifying and Reflecting on Limitations,* the more the participant can challenge their assumptions.

Specifying and Reflecting on Limitations may have been impacted by the Cognitive Overload phenomenon extracted with use of IPA. Participants expressed that the overload of information may have caused problems in their information processing abilities. Even though the intervention condition allowed then to control the speed of material presented, most participant mentioned this excess of information. Future use could emphasis use of pausing and rewinding to enable them to fully process information before moving on to the next stimuli.

Upon reflection participant 5 was quiet when asked to talk about the experience of using the material. At the time the researcher noted this and did not prompt the participant further as they seemed to have problem in communicating. This may have been a more radical case as experienced by other participants who were able to communicate how some elements of the material had too much information at once. It may be for participant 5 that much more of the content was too much information therefore they stopped engaging in the experiment due to overload. This may be explained by the inverted U theory (see Teigen, 1994), whereby performance is dictated by arousal, with peak arousal causing peak performance and low (such as boredom) or high (such as overload or panic) arousal cause lowered performance. Most participants experienced this high arousal via overload- and experientially significant enough for them to verbalise this afterwards. Participant 5 was a exemplar version in that they ceased interaction as they may have experienced discomfort from the experiment. (the researcher ensured they were ok before parting with the participant).

Further considerations based on the inverted U Theory should be implemented. These are four factors which affect how much pressure/information etc that an individual can take. These are skill level, personality, Trait anxiety, and Task completion. Although in the design of both conditional materials there was considerations on skill levels (targeted to one group of students), and Task complexity (reviewing content and providing control of videos), personality and Trait anxiety needs more considerations in future research. Indeed, current studies use inverted U Theory, also known as Yerkes-Dodson Law, to explain errors in domains with suggestion of more use in healthcare (Markose, Chiu, & Cousins, 2020). This highlights the importance of IPA, understanding of expert cognition, and other methods such as verbal protocol to investigate phenomenon in students to be powerful and in-depth enough to be able to measure these embedded phenomena within the raw data collected.

12.3 Effects of Heterogeneity before Intervention Upon Data Analysis

For example, the results for decision-making to uncover CAP processes were parallel with studies who also used VPA to uncover expert-novice differences when presented with the same scenarios. For example, Den Hartigh et al (2014) uncovered that more experienced players reported a substantial number of events that were away from the ball, with less skill players having a narrower range with the ball at the centre of their attention. Typical findings can present expert answers to appear counter intuitive (Dargar, Kennedy, Lai, Arikatla, & De, 2015). The CAPs in this study selected 1st rank choices that were ranked last by trainees. As acquisition of expertise occurs through use of such NTS training materials and real-world practice, knowledge and considerations of rules are eventually replaced by effortless pattern matching and retrieval of memories of previous experiences and actions. Although rationale for these choices can be hard to unpack, cognitive probing and written results helped to extract justifications. CAPs provided justifications with recognition of information which was paired with their schemata that had built their stronger mental models.

However, the results for the trainee participants' decision-making were not as expected as there were significantly different levels within the same cohort. This was not expected as the sample was from the same course year, and was assumed to infer a homogenous sample with both technical and NTS skills in positive correlation. Being in the same course level there was suggestion that these individuals were able to perform with a small deviation of change. If this deviation of change in performance was skewed with larger variance in one direction then an individual may be placed in a group which closer matches their typical performance levels. However, individuals such as mature students, can have decades of experience which may significantly affect their skills yet be placed in a group that does not match skill level due to the linearly fixed system of academic course progression. Additionally, there may be barriers to skill progression for some individuals that forms disparity of skill within the same group over time. Therefore, the demographic questions asked were included to establish the group had approximate technical skills similarities and experiences. For example, with an age range of 20-25 years it was impossible for 30 years' experience from an individual- experience generally correlates with expertise development. The technical skills from this homogenous sample were required to reduce effects from variability when performing exploration of their non-technical skills.

12.4 Decision-Making and Individual NTS Profiling

However, upon assessment of decision-making significant differences were suggested with verbal reports displaying many divergences of understanding for the same scenario, to produce the resultant ranking orders. Many (71%) had the same quantitative results in ranking but the cues and information used to form those ranking orders had substantial differences. Additionally, some participants showed already higher levels of decision making which may have suggested at least 14% of participant had advanced levels compared to their peers before intervention. This highlighted the individual differences which can confound measurements of non-technical skills and disable them from being defined as their counterpart correlating to technical skills (*see section 3.2)*. One attempt to address this could be use of a matched-pairs design. For example, participants 2 and 8 both displayed higher level NTS events and could potentially be matched in skill to better determine effects from each condition.

Yet, from exploring data in this study a limitation appeared within the NTS of participants which will also likely prevent matched pairs design in similar future research if decision-making is not homogenous before intervention. There was more individual difference than anticipated and caused difficulty in measuring and exploring NTS performance at a group level. Instead, the results were optimally displayed in a manner that was more aligned with individual profiles of their non-technical repertoire. There were common overarching events identified, such as staff members talking, or a surgeon asking for help. But reducing grain size of analysis by many levels past this as mentioned in section 9.5.1, produced content to appear dissimilar to an extent there was more suitability into producing and displaying 'personal profiles' of non-technical skill. Future research should proceed with caution when decision-making skills are heterogenous before intervention as this may cause participant's data to have dissimilarity that requires additional time and resource to solidify findings of a conditions' effects on cognition. No assistance from software is currently available to aid in analysis. Software can assess text for semantic meaning and syntactic structure, i.e. Stanford Part-of Speech tagger (Stanford, 2020). However, it will be several years before the ability to identify the type of non-technical category being used, based on techniques such as Critical Stylistics (*see Jefferies, 2009*).

12.5 Expert Input, Models of Expertise, and Data Analysis

There was greater dependence than appreciated on expert inclusion during the analysis of cognitive verbalizations, to explore specifically why participants interpreted cues differently. For example, when designing the experimental material in this study, the models used to consider how NTS skills could be transferred was with reference to models of Cognitive Apprenticeship and Surgical Coaching (*see 7.0.1.).* These models had frameworks into how constructive feedback can be provided while presenting cues that may be tacit if not pointed out by an expert. However, a more fitting framework should have been considered which could have provided more context from the CAPs to understand the 'personal profiles' reported. Future research must consider the model of *Intelligent cooperation* in the transfer of novice-expert information. Sutkin et al (2018) developed a novel framework of the Intelligent cooperation model for healthcare which involved assessment of sequences of coordinated exchanges from the surgical expert and the trainee. These exchanges can illuminate the needs of the trainee to learn efficiently, and enable the expert to form critical scaffolding feedback.

With implementation of this model, Sutkin, Littleton, and Kanter (2019) investigated videos and interviews of OT procedures whereby the surgical trainee was the primary operator. Their actions were guided by a surgeon who encouraged the desired behaviours by live feedback during the procedure. The results suggested that surgeons use several teaching mechanisms to speed up or slow down their trainees (*also evident for CAP2 see section 6.7.3*) depending on their need to enable improved skill development (slow down) or maintain patient safety status (speed up). If research applied this method to all staff in the OT it would allow improved understandings of the exact cues and moments in time when the expert improved the schema and scripts of the novice. Exploration of results would benefit as this precise measure of independent variables would allow better tracking of a novice's cognition as it deviated away from a desired process of NTS comprehension.

12.6 Formation of SAGAT Questions

Advantageously a mixed measures design incorporating quantitative measures can aid in understanding if such issues occur, as trends can be easier identified with standardized analytical processes. The SAGAT questions were an example of this. One way the information from the task analysis was reused was the aid in the creation of the ordinal data of the SAGAT questions. The final questions used in the SAGAT probing were included due to the cognitive interviews performed with the 3 CAPs participating in this study as well as mostly being linked to parallel research. The questions asked (*see Appendix Figure 7 and Figures 8A-D*) elicited information that was frequently used to explain how they maintained a high-performance level in theatre. For a researcher to form a domain-specific SAGAT, the task-critical SA cues are firstly needed. These cues comprise of the information in a given environment that facilitates the perceptual anticipation and projection of future situational state(s).

An example of this was the use of music distraction as a question. It was through a combination of interviews and academic research into communication related surgical incidents that this was added to the SAGAT. Music has been shown to have many benefits on performance (e.g. Oomens, Kleinrensink, & Jeekel, 2019), but it also can disrupt flow (El Bardissi, Wiegmann, Dearani, Daly, & Sundt, 2008), and be detrimental when unpredicted critical events occur (MacDonald & Schlesinger, 2018). To use the SAGAT in the context of the current study there was requirement to identify appropriate cues in the Operating Theatre. The benefit of virtual simulation was the ability to add music and control its volume -there was no music during recording but was added due to its role in NTS related errors.

From this, although the SAGAT scores showed that 71% of all participants correctly identified the increase in music loudness the cognitive verbalizations from the experimental condition suggested their understanding of its importance increased to become a cue that had enough value for acknowledgement, in relation to error prevention. For example, participant 2 stated, '*has the music gone up? Maybe it's just me paying more attention to things but it sounds like its more distracting'*(l16). Different to many other SAGAT questions this was not a visual cue and not apparent unless the participant had a predisposition to attend, or has cause to attend to the cue. This was one question which supported the validity of the SAGAT when providing context to measuring the changes of NTS towards patient safety.

Another example of the rationale for SAGAT questions was the inclusion of a time-based question - '*As accurate as possible, how many seconds was the scenario?*'. An important factor in situational awareness is time; both the perception of time passing and the objective temporal measure of event continuation. The control condition’s intervention material did improve participants’ accuracy towards the correct answer (p=0.027), but was not as accurate as the experimental condition due to a less accurate mean pre-intervention score. As there was a small amount of data along with a marginal significance value and effect size (p=0.054) to further investigate this difference may produce incorrect assumptions on the effect of conditions. However, this exploration did support the value in choosing SAGAT questions that have real world meaning when data is analysed. For this study, the effects of condition upon the distortion of time perception has importance for many roles in individuals’ Situational Awareness abilities (Endsley & Robertson, 1996; Gillespie, Gwinner, Fairweather, & Chaboyer, 2013) and the CAPs provided support for the role of time awareness on the ability to predict and address undesired events (*see section 6.7.3).* Little research has been available on Time in VR without reference to motion or intended manipulation from virtual independent variables (such as a Sun). However generally there has been suggestion that an array of virtual material can influence time perception. By either providing perception of compresses time during chemotherapy (Schneider, Kisby, & Flint, 2011), or by time acceleration/overestimation (Schatzschneider, Bruder, & Steinicke, 2016*).* The result from this thesis could suggest minimal effect on time perception with VR headset usage. Therefore, training without the repercussion of weakened time measurement abilities would prevent reduced skills in time-critical and time-based predictive events.

12.7 Cohens kappa Usage

Like the formation of the SAGAT questions, another methodological strength was the use of Cohen’s kappa. Certainly, percentage agreement was not used as it has been frequently criticized for its ability to inflate the measure of agreement as it does not account for the likelihood by change for such agreement to occur. Percentage agreement for IRR is a quicker and weaker method that has susceptibility in the failure to account for chance (McHugh, 2012). From 0 to 1 (no agreement to perfect agreement) simple calculations can find the percentage of a pair of coders on their coding agreement. Additionally, categories may be added or modified to allow easier agreement which then would not hold the main focus of in-depth analysis of the data but being methodologically ‘safe’. If there were 2 categories to encode into there is 50% chance of IRR agreement. However, the more categories that are used the chance if IRR agreement then reduces. Conversely, too many categories may cause overlap of categories in certain contexts and effect reliability. In the context of this thesis, there were 4 overarching NTS categories at the highest level that were broken down into 21 separate categories of non-technical skills. This had the same setup as Mitchell et al (2011) who also used two inter-raters to maintain k>.81 from the same 4 overarching NTS categories and a k>.73 from their separately defined bottom-level categories. Additionally, Walker et al (2001) gauged interrater reliability for both raters at 0.9 and 0.7 respectively for an on-road vehicle investigation. This suggested that if random allocation occurred, Cohens kappa would be approximately 0.25 for overarching NTS categories with less reliability when these were broken down. Yet, k=.8 in this thesis and similarly suggested that with the inclusion chance as a factor there was still a moderate/strong IRR gained from the coding setup and checks performed.

This result was gained by providing the 2nd inter-rater a 6-page booklet on the theories, models, and terminologies involved along with an hour’s discussion on the data and coding processes. The goal of this was to increase the shared meaning of each theme. The coding scheme used in this study was theoretically driven but ultimately theme based therefore the kappa result could be viewed as a reflection of shared meaning accuracy of themes. Certainly, there was one meeting to confer/discuss how and why differences in coding had formed. Initially, the coding scheme terminology was not clear enough for the 2nd rater and was revised. This increased similarity after the clarification and the primary coder recheck and improved categorization accuracy because of this.

12.8 Introspection and Transfer of Skill to OT

The effects on participants' own introspection was not accounted for and produced a negative effect when an occurrence was observed. Participant 13 stated, '*I feel not as good as I thought I was, I don't prepare tools or anything I just get asked* (I5)*'.* The impact of this was not followed up and future research should put measures in place as participants may have introspection-based concern into their strengths and weakness which may impact them in the OT. Being the transfer of effect from the experiment to real-life, this also questioned other factors which may have changed the participant's views on their, and colleagues NTS because of both condition's intervention material. Initially, a suggestion to investigate and/or prevent this may be to use real-life scenarios for the pre-and post- intervention measures. There may be attraction for future research to setup simulation in a real OT with surgical actors and a highly controlled script/scenario. Unfortunately, a method that allowed multiple observers to talk aloud without interfering with the simulation would be difficult to implement. This would result in many viewpoints of the OT with cues in different proximity to participants and high chance of interreference with other participant's out loud cognitive verbalizations. Any effects occurring such as introspection would be better investigated by follow up observations in the OT of select participants to ascertain if their viewpoints had changed.

12.9 Future Directions and Conclusion

There should be more support for the examination of efficacy and the distribution of digital material with evidence to improve NTS and cognitive processes. If a realistic deployment can be found for VR usage to be a synergy from other training materials, then more accessibility would mean NTS related training would increase. For example, such VR NTS training material could be provided to individuals prior to attending a NOTSS course. This would allow them to get an initial sense of NTS, and allow efficient delivery during the NOTSS workshops along with a standardised material to compare if the user's session were recorded on their VR device (such as in this thesis). Previous acknowledgement that valid and reliable assessment of NTS was required throughout training to facilitate learning was one reason why the NOTSS was created, and the same goal was attempted in this study.

Moreover, when issues may prevent such lectures and workshops from occurring, a possible combination of- or a 'best of both worlds' - training material would allow accessible practice that can guide users to improve their NTS abilities. Common issues such as the costs to setup workshops and bootcamps along with restrictions in times and locations may be alleviated with a combination of this 24/7-hour accessible material. Indeed, with less common issues such as the recent COVID-19 Coronavirus outbreak (Zhu et al, 2020), governments and health authorities around the world had created guidelines and restrictions which reduced the ability to travel for months at a time. Such downloadable NTS training to low cost headsets would allow users to increase their representations of an array of medical and surgical scenarios and improve abilities as significantly occurring in this thesis. Such effortful and deliberate processes are consistencies required in the formation of schema orientated towards expertise.

In conclusion, there were no significant increase in NTS abilities due to the control condition material compared to the experimental condition when comparing quantitative results. However, there were significant increases in participant's abilities to transform information and attempt to align their mental models with the staff in the scenario after intervention. Qualitative exploration suggested increased cue scanning and attentiveness to socially-orientated NTS cues for the experimental condition however the individual variability in the interpretations used on information hindered the ability to explore changes at a group level. Participant's reflections (*see Figure 18*) could be captured and analysed to concur with any findings and further. Additionally, several sessions followed by observations in the OT could further explore benefits and limitation of the different training contents along with evidence of long-term transfer of skills. These were outside the scope of this thesis.

References

Aebersold, M., Rasmussen, J., & Mulrenin, T. (2020). Virtual Everest: Immersive Virtual Reality Can Improve the Simulation Experience. *Clinical Simulation in Nursing*, *38*, 1-4.

Afkari, H., Bednarik, R., Makala, S., & Eivazi, S. (2016). Mechanisms for maintaining situation awareness in the micro-neurosurgical operating room. *International Journal of Human-Computer Studies*, *95*, 1-14.

Agha, R. A., & Fowler, A. J. (2015). The role and validity of surgical simulation. *International surgery*, *100*(2), 350-357.

Ahmed, S., & Dann, S. (2015). The virtual future of education and training. *The Bulletin of the Royal College of Surgeons of England*, *97*(10), 431-432.

Aiken, M. P., & Berry, M. J. (2015). Posttraumatic stress disorder: possibilities for olfaction and virtual reality exposure therapy. *Virtual Reality*, 1-15.

Aïm, F., Lonjon, G., Hannouche, D., & Nizard, R. (2016). Effectiveness of Virtual Reality Training in Orthopaedic Surgery. *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, 32(1), 224-232.

Alaker, M., Wynn, G. R., & Arulampalam, T. (2016). Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *International Journal of Surgery*, *29*, 85-94.

Alsmith, A. (2015). Perspectival Structure and Vestibular Processing. *Open MIND. Frankfurt a. M., GER: MIND Group*.

Aman, M., Kistler, W. D., Barrington, Q., Sharp, W., Persky, S., & Shaw, P. (2020). Using virtual reality to define the mechanisms linking symptoms with cognitive deficits in attention deficit hyperactivity disorder. *Scientific Reports (Nature Publisher Group)*, *10*(1).

Andersen, S. A. W., Mikkelsen, P. T., Konge, L., Cayé-Thomasen, P., & Sørensen, M. S. (2016). The effect of implementing cognitive load theory-based design principles in virtual reality simulation training of surgical skills: a randomized controlled trial. *Advances in Simulation*, *1*(1), 20.

Andersen, S. A. W., Mikkelsen, P. T., Konge, L., Cayé-Thomasen, P., & Sørensen, M. S. (2016). Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation. *The Laryngoscope*, *126*(2), E74-E79.

Anderson, O., Davis, R., Hanna, G. B., & Vincent, C. A. (2013). Surgical adverse events: a systematic review. *The American Journal of Surgery*, *206*(2), 253-262.

AO Foundation- Midface Le Fort I. (2019). Aofoundation.org. Retrieved November 12, 2019, from https://www2.aofoundation.org/wps/portal/surgery?bone=CMF&segment=Midface&classification=92-Le Fort, Type I, Linear fracture&showPage=indication.

Art Institutes (2020). Games design software you can use to make your own games. Available from: https://www.artinstitutes.edu/about/blog/aig-3-game-design-software-tools-you-can-use-to-make-your-own-games. Last Accessed: 20 Feb 2020.

Ahmed, K., Aydin, A., Dasgupta, P., Khan, M. S., & McCabe, J. E. (2015). A novel cadaveric simulation program in urology. *Journal of surgical education*, *72*(4), 556-565.

Amalberti, R., & Deblon, F. (1992). Cognitive modelling of fighter aircraft process control: a step towards an intelligent on-board assistance system. *International Journal of Man-Machine Studies*, *36*(5), 639-671.

Armstrong, C. M., Reger, G. M., Edwards, J., Rizzo, A. A., Courtney, C. G., & Parsons, T. D. (2013). Validity of the Virtual Reality Stroop Task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology*, *35*(2), 113-123.

Awad, S. S., Fagan, S. P., Bellows, C., Albo, D., Green-Rashad, B., De La Garza, M., & Berger, D. H. (2005). Bridging the communication gap in the operating room with medical team training. *The American Journal of Surgery*, *190*(5), 770-774.

Ayoub, A., Khambay, B., Benington, P., Green, L., Moos, K., & Walker, F. (2013). *Handbook of Orthognathic Treatment: A Team Approach*. John Wiley & Sons.

Bahrami, P., Graham, S. J., Grantcharov, T. P., Cusimano, M. D., Rotstein, O. D., Mansur, A., & Schweizer, T. A. (2014). Neuroanatomical correlates of laparoscopic surgery training. *Surgical Endoscopy*, 28(7), 2189-2198.

Bainbridge, L., & Sanderson, P. (1990). Verbal protocol analysis. In: Wilson, J.R., Corlett, E.N. (Eds.), Evaluation of Human Work: A Practical Ergonomics Methodology, second ed. Taylor & Francis, Ltd, London, UK, pp. 169–201.

Balch, C. M., Shanafelt, T. D., Bechamps, G., Russell, T., Dyrbye, L., Satele, D., & Freischlag, J. (2010). Burnout and medical errors among American surgeons. *Annals of surgery*, *251*(6), 995-1000.

Banerjee, A., & Gavade, S. (2015). Integration of Devices in Operation Room for Reduction in Cognitive Load during Surgical Procedures. In *2015 International Conference on Computational Science and Computational Intelligence (CSCI)*. pp. 821-824. IEEE.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). What the drivers do and do not tell you: using verbal protocol analysis to investigate driver behaviour in emergency situations. *Ergonomics*, *57*(3), 332-342.

Bass, B. M. (1990). From transactional to transformational leadership: Learning to share the vision. *Organizational dynamics*, *18*(3), 19-31.

Borracci, R. A., & Arribalzaga, E. B. (2018). The incidence of overconfidence and underconfidence effects in medical student examinations. *Journal of surgical education*, *75*(5), 1223-1229.

Berridge, C., Jain, S., & Biyani, C. S. (2019). Current and future simulation in urological surgery training. *Trends in Urology & Men's Health*, *10*(3), 16-18.

Bethune, R., Sasirekha, G., Sahu, A., Cawthorn, S., & Pullyblank, A. (2011). Use of briefings and debriefings as a tool in improving teamwork, efficiency, and communication in the operating theatre. *Postgraduate medical journal*, *87*(1027), 331-334.

Birkmeyer, J. D., Finks, J. F., O'Reilly, A., Oerline, M., Carlin, A. M., Nunn, A. R., ... & Birkmeyer, N. J. (2013). Surgical skill and complication rates after bariatric surgery. *New England Journal of Medicine*, *369*(15), 1434-1442.

Black, S. A., Nestel, D. F., Kneebone, R. L., & Wolfe, J. H. N. (2010). Assessment of surgical competence at carotid endarterectomy under local anaesthesia in a simulated operating theatre. *British journal of surgery*, *97*(4), 511-516.

Blackmur, J. P., Clement, R. G., Brady, R. R., & Oliver, C. W. (2013). Surgical training 2.0: How contemporary developments in information technology can augment surgical training. *the surgeon*, *11*(2), 105-112.

Bommer, C., Sullivan, S., Campbell, K., Ahola, Z., Agarwal, S., O'Rourke, A., ... & Liepert, A. E. (2018). Pre-simulation orientation for medical trainees: An approach to decrease anxiety and improve confidence and performance. *The American Journal of Surgery*, *215*(2), 266-271.

Bonrath, E. M., Dedy, N. J., Gordon, L. E., & Grantcharov, T. P. (2015). Comprehensive surgical coaching enhances surgical skill in the operating room. *Annals of surgery*, *262*(2), 205-212.

Bordnick, P. S., Traylor, A. C., Carter, B. L., & Graap, K. M. (2012). A feasibility study of virtual reality-based coping skills training for nicotine dependence. *Research on social work practice*, *22*(3), 293-300.

Bourrier, Y., Jambon, F., Garbay, C., & Luengo, V. (2016). An Interaction Centred Approach to the teaching of Non-technical Skills in a Virtual Environment. Computer Laboratory of Paris 6.

Botden, S. M., Torab, F., Buzink, S. N., & Jakimowicz, J. J. (2008). The importance of haptic feedback in laparoscopic suturing training and the additive value of virtual reality simulation. *Surgical endoscopy*, 22(5), 1214-1222.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77. doi:10.1191/1478088706qp063oa

Braun, V., and Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. London: Sage.

Braun, V., Clarke, V., and Terry, G. (2015). Thematic analysis. In P. Rohleder and A. Lyons (Eds.), Qualitative research in clinical and health psychology (pp. 95–113). Basingstoke: Palgrave MacMillan.

Brennan, S. (2006). NOTSS course (non-technical skills for surgeons). *BMJ*, *333*(7567), s98-s98.

Brunckhorst, O., Shahid, S., Aydin, A., Khan, S., McIlhenny, C., Brewin, J., ... & Dasgupta, P. (2015). The relationship between technical and nontechnical skills within a simulation-based ureteroscopy training environment. *Journal of surgical education*, *72*(5), 1039-1044.

British Psychological Society (2018). Code of Ethics and Conduct. [online] Leicester: The British Psychological Society. Available at: https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Policy%20-%20Files/BPS%20Code%20of%20Ethics%20and%20Conduct%20%28Updated%20July%202018%29.pdf [Accessed 23 May 2019].

Buckley, C. E., Kavanagh, D. O., Nugent, E., Ryan, D., Traynor, O. J., & Neary, P. C. (2014). The impact of aptitude on the learning curve for laparoscopic suturing. *The American Journal of Surgery*, *207*(2), 263-270.

Burton, Z. A., Guerreiro, F., Turner, M., & Hackett, R. (2018). Mad as a hatter? Evaluating doctors' recall of names in theatres and attitudes towards adopting# theatrecapchallenge. *British journal of anaesthesia*, *121*(4), 984-986.

Brunckhorst, O., Aydin, A., Abboudi, H., Sahai, A., Khan, M. S., Dasgupta, P., & Ahmed, K. (2015). Simulation-based ureteroscopy training: a systematic review. *Journal of surgical education*, *72*(1), 135-143.

Cadet, L. B., & Chainay, H. (2020). Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human-Computer Studies*, *144*, 102506.

Campbell, J., Tirapelle, L., Yates, K., Clark, R., Inaba, K., Green, D., & Sullivan, M. (2011). The effectiveness of a cognitive task analysis informed curriculum to increase self-efficacy and improve performance for an open cricothyrotomy. *Journal of surgical education*, 68(5), 403-407Caro, P. W. (1988). Flight training and simulation. In *Human factors in aviation* (pp. 229-261). Academic Press.

Chapman, S. J., Hakeem, A. R., Marangoni, G., & Prasad, K. R. (2013). How can we enhance undergraduate medical training in the operating room? A survey of student attitudes and opinions. *Journal of surgical education*, *70*(3), 326-333.

Chapman, G. B., & Sonnenberg, F. A. (2003). *Decision making in health care: theory, psychology, and applications*. Cambridge University Press.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J. P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis & Prevention*, *59*, 26-37.

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The journal of the learning sciences*, *6*(3), 271-315.

Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (Eds.). (2000). *Cognitive task analysis*. Psychology Press.

Choudhury, N., Gélinas-Phaneuf, N., Delorme, S., & Del Maestro, R. (2013). Fundamentals of neurosurgery: virtual reality tasks for training and evaluation of technical skills. *World neurosurgery*, *80*(5).

Christian, C. K., Gustafson, M. L., Roth, E. M., Sheridan, T. B., Gandhi, T. K., Dwyer, K., ... & Dierks, M. M. (2006). A prospective study of patient safety in the operating room. *Surgery*, *139*(2), 159-173.

Clark, R. E., Feldon, D., van Merriënboer, J. J. G., Yates, K., and Early, S. (2008). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.). Handbook of research on educational communications and technology (3rd ed.) (pp. 577-593). New York: Macmillan/Gale.

Clark, R. E., Yates, K., Early, S., Moulton, K., Silber, K. H., & Foshay, R. (2010). An analysis of the failure of electronic media and discovery-based learning: Evidence for the performance benefits of guided training methods. *Handbook of training and improving workplace performance*, *1*, 263-297.

Coffey, J. W., & Hoffman, R. R. (2003). Knowledge modelling for the preservation of institutional memory. *Journal of Knowledge Management*, *7*(3), 38-52.

Coffey-Zern, S., Calvi, R. M., Vorrasi, J. S., & Meara, D. J. (2015). Incorporating simulation into oral and maxillofacial surgery residency education and training: christiana care's method. *Journal of Oral and Maxillofacial Surgery*, *73*(7), 1244-1245.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), 37-46. Available from: http://www.ric.edu/faculty/organic/coge/cohen1960.pdf. Last Accessed: 22 Feb 2020.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American educator*, *15*(3), 6-11.

Cologne, K. G., Zehetner, J., Liwanag, L., Cash, C., Senagore, A. J., & Lipham, J. C. (2015). Three-dimensional laparoscopy: does improved visualization decrease the learning curve among trainees in advanced procedures?. *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, *25*(4), 321-323.

Cook, D. A., Erwin, P. J., & Triola, M. M. (2010). Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Academic Medicine*, *85*(10), 1589-1602.

Cooper, J. B., Newbower, R. S., Long, C. D., & McPeek, B. (2002). Preventable anaesthesia mishaps: a study of human factors. *Quality and Safety in Health Care*, *11*(3), 277-282.

Crebbin, W., Beasley, S. W., & Watters, D. A. (2013). Clinical decision-making: how surgeons do it. *ANZ journal of surgery*, *83*(6), 422-428.

Crichton, M., O'Connor, P., & Flin, R. (2013). *Safety at the sharp end: a guide to non-technical skills*. Ashgate Publishing, Ltd.

Crispen, P., & Hoffman, R. R. (2016). How Many Experts?. *IEEE Intelligent Systems*, *31*(6), 56-62.

Cristancho, S. M., Apramian, T., Vanstone, M., Lingard, L., Ott, M., Forbes, T., & Novick, R. (2016). Thinking like an expert: surgical decision making as a cyclical process of being aware. *The American Journal of Surgery*, *211*(1), 64-69.

Crutchfield, D. N., Moertl, P. M., & Ohrt, D. Expertise and Chess: A Pilot Study Comparing Situation Awareness Methodologies.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272-309.

Cuschieri, A., Francis, N., Crosby, J., & Hanna, G. B. (2001). What do master surgeons think of surgical competence and revalidation? *The American journal of surgery*, 182(2), 110-116.

Dankelman, J., & Grimbergen, C. A. (2005). Systems approach to reduce errors in surgery. *Surgical endoscopy and other interventional techniques*, *19*(8), 1017-1021.

Dargar, S., Kennedy, R., Lai, W., Arikatla, V., & De, S. (2015). Towards immersive virtual reality (iVR): a route to surgical expertise. *Journal of computational surgery*, *2*(1), 2.

Davincisurgery.com (2020). DaVinci- Intuitive for Patients. Intuitive. Retrieved 02 January 2019 from https://www.davincisurgery.com/.

Davis, R. E., Sevdalis, N., Jacklin, R., & Vincent, C. A. (2012). An examination of opportunities for the active patient in improving patient safety. *Journal of Patient Safety*, *8*(1), 36-43.

Demirel, D., Butler, K. L., Halic, T., Sankaranarayanan, G., Spindler, D., Cao, C., ... & deMoya, M. A. (2016). A hierarchical task analysis of cricothyroidotomy procedure for a virtual airway skills trainer simulator. *The American Journal of Surgery*, *212*(3), 475-484.

Den Hartigh, R. J., Van Der Steen, S., De Meij, M., Van Yperen, N. W., Gernigon, C., & Van Geert, P. L. (2014). Characterising expert representations during real-time action: A Skill Theory application to soccer. *Journal of Cognitive Psychology*, *26*(7), 754-767.

Dias, R. D., Ngo-Howard, M. C., Boskovski, M. T., Zenati, M. A., & Yule, S. J. (2018). Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. *British Journal of Surgery*, *105*(5), 491-501.

Diemer, J., Mühlberger, A., Pauli, P., & Zwanzger, P. (2014). Virtual reality exposure in anxiety disorders: Impact on psychophysiological reactivity. *The World Journal of Biological Psychiatry*, *15*(6), 427-442.

DiLullo, C., McGee, P., & Kriebel, R. M. (2011). Demystifying the Millennial student: A reassessment in measures of character and engagement in professional education. *Anatomical sciences education*, *4*(4), 214-226.

Drew, T., Võ, M. L. H., & Wolfe, J. M. (2013). The invisible gorilla strikes again sustained inattentional blindness in expert observers. *Psychological Science*, 24(9), 1848-1853.

Dreyfus, H. L., & Dreyfus, S. E. (1996). The relationship of theory and practice in the acquisition of skill. *Expertise in nursing practice: Caring, clinical judgment, and ethics*, 29-47. In Benner, P., Benner, P. E., Tanner, C. A., & Chesla, C. A. (2009). *Expertise in nursing practice: Caring, clinical judgment, and ethics*. Chapter 2. Springer Publishing Company. Available from: https://books.google.co.uk/books?id=wzxDZXJShhkC&dq=The+relationship+of+theory+and+practice+in+the+acquisition+of+skill&source=gbs_navlinks_s. Last Accessed: 2 Feb 2020

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition* (No. ORC-80-2). California University Berkeley Operations Research Center.

Durning, S., Artino Jr, A. R., Pangaro, L., van der Vleuten, C. P., & Schuwirth, L. (2011). Context and clinical reasoning: understanding the perspective of the expert's voice. *Medical education*, 45(9), 927-938.

Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., & Manning, C. A. (1998). En route operational errors and situational awareness. *The International Journal of Aviation Psychology*, *8*(2), 177-194.

Eason, A. (2019). *Applying 'spidey Sense': Emotional Intelligence and Police Decision-making at Domestic Incidents* (Doctoral dissertation, University of Portsmouth).

ElBardissi, A. W., Wiegmann, D. A., Henrickson, S., Wadhera, R., & Sundt III, T. M. (2008). Identifying methods to improve heart surgery: an operative approach and strategy for implementation on an organizational level. *European Journal of Cardio-thoracic Surgery*, *34*(5), 1027-1033.

Elledge, R., McAleer, S., Thakar, M., Begum, F., Singhota, S., & Grew, N. (2016). Use of a virtual learning environment for training in maxillofacial emergencies: impact on the knowledge and attitudes of staff in accident and emergency departments. *British Journal of Oral and Maxillofacial Surgery*, *54*(2), 166-169.

El-Shafy, I. A., Delgado, J., Akerman, M., Bullaro, F., Christopherson, N. A., & Prince, J. M. (2018). Closed-loop communication improves task completion in paediatric trauma resuscitation. *Journal of surgical education*, *75*(1), 58-64.

Endsley, M. R. (1988, May). Situation awareness global assessment technique (SAGAT). In *Proceedings of the IEEE 1988 national aerospace and electronics conference* (pp. 789-795). IEEE.

Endsley, M. R., Bolte, B., & Jones, G.J. (2003). Designing for Situational Awareness- An Approach to User-Centered Design. Taylor and Francis: CRC Press, NY.

Endsley, M. R., & Garland, D. J. (2000). Theoretical underpinnings of situation awareness: A critical review. *Situation awareness analysis and measurement*, *1*, 24.

Endsley, M. R. (2016). *Designing for situation awareness: An approach to user-centred design*. CRC press.

Endsley, M. R., Bolstad, C. A., Jones, D. G., & Riley, J. M. (2003). Situation Awareness oriented design: from user's cognitive requirements to creating effective supporting technologies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, No. 3, pp. 268-272). Sage CA: Los Angeles, CA: SAGE Publications.

Endsley, M. R., & Robertson, M. M. (1996). Team situation awareness in aviation maintenance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 40, 21, 1077-1081). SAGE Publications.

Endsley, M. R., Sollenberger, R., Nakata, A., & Stein, E. S. (2000). *Situation awareness in air traffic control: Enhanced displays for advanced operations*. Federal Aviation Administration technical centre, Atlantic City. NJ.

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British journal of educational psychology.* 70(1):113–36

Ericsson, K. A. (2018). Capturing Expert Thought with Protocol Analysis: Concurrent Verbalizations of Thinking during Experts' Performance on Representative Tasks. *The Cambridge Handbook of Expertise and Expert Performance*, 192.

Ericsson, K. A., Hoffman, R. R., & Kozbelt, A. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

Ericsson KA, Lehmann AC. (1996). Expert and exceptional performance: evidence on maximal adaptations on task constraints. Annual Review of Psychology. 47:273-305.

Ericsson, K. A., & Moxley, J. H. (2019). Thinking aloud during superior performance on tasks involving decision making. *A Handbook of Process Tracing Methods*, 286.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard business review*, *85*(7/8), 114.

Ericsson, K. A., & Simon, H. A. (1993). Protocol Analysis: Verbal reports as data, the MIT press.Ericsson, K. A., & Simon, H. A. (1998). Protocol analysis. *A companion to cognitive science*, *14*, 425-432.

Eurocontrol. (2020). Human Performance Repository. Available from: https://ext.eurocontrol.int/ehp/?q=node/1601. Last Accessed: 2 Feb 2020.

Evans, A.W. (2001). Assessing competence in surgical dentistry. *British Dental Journal* 190, 7: 343–346.

Fabri, P. J., & Zayas-Castro, J. L. (2008). Human error, not communication and systems, underlies surgical complications. Surgery, 144(4), 557-565.

Farag, C., Troiani, V., Bonner, M., Powers, C., Avants, B., Gee, J., & Grossman, M. (2010). Hierarchical organization of scripts: converging evidence from FMRI and frontotemporal degeneration. *Cerebral Cortex*, *20*(10), 2453-2463.

Felnhofer A, Kothgassner OD, Hetterle T, Beutl L, Hlavacs H, Kryspin-Exner, I. (2015). Afraid to Be There? Evaluating the Relation Between Presence, Self-Reported Anxiety, and Heart Rate in a Virtual Public Speaking Task. *Cyberpsychology, Behaviour, and Social Networking.* 17(5):310–316. doi: 10.1089/cyber.2013.0472.

Felnhofer, A., Kothgassner, O. D., Schmidt, M., Heinzle, A. K., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2015). Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International journal of human-computer studies*, *82*, 48-56.

Feussner, H. (2017). Surgery 4.0. In: Thuemmler C, Bai C (eds). *Health 4.0: How Virtualization and Big Data are Revolutionizing healthcare*. Springer, Champion, pp 91–107.

Fiedler FE. A *theory of leadership effectiveness*. New York: McGraw-Hill, 1967.

Fitts, P.M., & Jones, R.E. (1947). Analysis *of factors contributing to 460 "pilot error" experiences in operating aircraft controls* (Report No. TSEAA-694-12). In *Human Factors in Air Transport: Understanding Behaviour and Performance in Aviation* (pp. 1-10). Retrieved from: https://books.google.co.uk/books?id=qPGrDwAAQBAJ&printsec=frontcover#v=onepage&q&f=false

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fletcher, G. C. L., McGeorge, P., Flin, R. H., Glavin, R. J., & Maran, N. J. (2002). The role of non-technical skills in anaesthesia: a review of current literature. *British Journal of Anaesthesia*, *88*(3), 418-429.

Flin, R., Fletcher, G., Glavin, R., Maran, N. & Patey, R. (2003). *Anaesthetists' Non-Technical Skills (ANTS) System Handbook*. Retrieved from http://www.abdn.ac.uk/iprc/ants. Last Accessed: 20/06/2019.

Flin, R., Martin, L., Goeters, K. M., Hormann, H. J., Amalberti, R., Valot, C., & Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Human Factors and Aerospace Safety*, *3*, 97-120.

Flin, R. & Mitchell, L. (2008). Non-technical skills of the operating theatre scrub nurse: literature review. *Journal of advanced nursing*, *63*(1), 15-24.

Flin, R., O'Connor, P. & Crichton, M. (2007) *Safety at the Sharp End: A Guide to Non-Technical Skills.* Aldershot: Ashgate.

Flin, R., O'Connor, P., & Mearns, K. (2002). Crew resource management: improving teamwork in high reliability industries. *Team performance management: an international journal*, *8*(3/4), 68-78.

Flin, R., Youngson G.G. & Yule S. (2007) How do surgeons make intra-operative decisions. *Quality and Safety in Healthcare* 16, 235–239.

Flin, R., Youngson, G. G., & Yule, S. (Eds.). (2015). *Enhancing Surgical Performance: A Primer in Non-Technical Skills*. CRC Press.

Flin, R., Yule, S., Paterson-Brown, S., Maran, N., Rowley, D., & Youngson, G. (2006). Experimental Evaluation of a Behavioural Marker System for Surgeons' Non-Technical Skills (NOTSS). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, No. 10, pp. 969-972). Sage CA: Los Angeles, CA: SAGE Publications.

Flin, R., Yule, S., Paterson-Brown, S., Maran, N., Rowley, D., & Youngson, G. (2007). Teaching surgeons about non-technical skills. *The Surgeon*, *5*(2), 86-89.

Flowers, G., and Aggarwal, R., (2014). Second Life: a novel simulation platform for the training of surgical residents. *Expert review of medical devices* 11, 2: 101–103.

Forbes, T.W. (1939). The normal automobile driver as a traffic problem. The *Journal of General Psychology, 20,* 471-474.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, *32*(200), 675-701.

Gabor, E. (2015). Words Matter: radio misunderstandings in wildland firefighting. *International Journal of Wildland Fire*, *24*(4), 580-588.

Gahm, G., Reger, G., Ingram, M. V., Reger, M., & Rizzo, A. (2015). A Multisite, Randomized Clinical Trial of Virtual Reality and Prolonged Exposure Therapy for Active Duty Soldiers with PTSD. *Geneva foundation Tacoma* WA.

Garg, A. X., Norman, G., & Sperotable, L. (2001). How medical students learn spatial anatomy. *The Lancet*, *357*(9253), 363-364.

Gawande, A. A., Thomas, E. J., Zinner, M. J., & Brennan, T. A. (1999). The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*, *126*(1), 66-75.

Gawande, A. A., Zinner, M. J., Studdert, D. M., & Brennan, T. A. (2003). Analysis of errors reported by surgeons at three teaching hospitals. *Surgery*, *133*(6), 614-621.

Gerstle, C. R. (2018). Parallels in safety between aviation and healthcare. *Journal of pediatric surgery*, *53*(5), 875-878.

Gjeraa, K., Dieckmann, P., Spanager, L., Petersen, R. H., Østergaard, D., Park, Y. S., & Konge, L. (2019). Exploring Shared Mental Models of Surgical Teams in Video-Assisted Thoracoscopic Surgery Lobectomy. *The Annals of thoracic surgery*, *107*(3), 954-961.

Glotzbach-Schoon, E., Andreatta, M., Reif, A., Ewald, H., Tröger, C., Baumann, C., ... & Pauli, P. (2013). Contextual fear conditioning in virtual reality is affected by 5HTTLPR and NPSR1 polymorphisms: effects on fear-potentiated startle. *Frontiers in behavioural neuroscience*, *7*, 31.

Goodson, A. M., Payne, K. F., Tahim, A., Cabot, L., & Fan, K. (2013). Awareness of oral and maxillofacial surgery as a specialty and potential career pathway amongst UK medical undergraduates. *The Surgeon*, *11*(2), 92-95.

Gordon, M., Darbyshire, D., Baker P. (2012). Non-technical skills training to enhance patient safety: a systematic review. Medical Education. 46:1042–54.

Gore, J., Flin, R., Stanton, N., & Wong, W. (Eds.). (2015). *Applications of Naturalistic Decision Making*. British Psychological Society.

Gore, J., Ward, P., Conway, G. E., Ormerod, T. C., Wong, B. W., & Stanton, N. A. (2018). Naturalistic decision making: navigating uncertainty in complex sociotechnical work. p521-524.

Guru, K. A., Esfahani, E. T., Raza, S. J., Bhat, R., Wang, K., Hammond, Y., ... & Chowriappa, A. J. (2015). Cognitive skills assessment during robot-assisted surgery: separating the wheat from the chaff. *BJU international*, *115*(1), 166-174.

Graafland, M., Schraagen, J. M. C., Boermeester, M. A., Bemelman, W. A., & Schijven, M. P. (2015). Training situational awareness to reduce surgical errors in the operating room. *British journal of surgery*, *102*(1), 16-23.

Gillespie, B. M., Gwinner, K., Fairweather, N., & Chaboyer, W. (2013). Building shared situational awareness in surgery through distributed dialog. *Journal of multidisciplinary healthcare*, *6*, 109.

Grand View Research, (2019). Surgical Instruments Industry Report. Retrieved July 1, 2019, from https://www.grandviewresearch.com/industry-analysis/surgical-equipment-market

Green, B., Tsiroyannis, C., & Brennan, P. A. (2016). Human factors–recognising and minimising errors in our day to day practice. *Oral diseases*, *22*(1), 19-22

Greenberg, C. C., & Klingensmith, M. E. (2015). The continuum of coaching: opportunities for surgical improvement at all levels. *Annals of surgery*, *262*(2), 217. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558915/

Gronlund, S. D., Dougherty, M. R., Durso, F. T., & Canning, J. M. (2001). *Planning in air traffic control* (No. DOT/FAA/AM-01/16). United States. Office of Aerospace Medicine.

Gross, B., Rusin, L., Kiesewetter, J., Zottmann, J. M., Fischer, M. R., Prückner, S., & Zech, A. (2019). Crew resource management training in healthcare: a systematic review of intervention design, training conditions and evaluation. *BMJ open*, *9*(2), retrieved from: https://bmjopen.bmj.com/content/bmjopen/9/2/e025247.full.pdf

Haggerty, C. J., & Laughlin, R. M. (2015). *Operative Atlas of Oral and Maxillofacial Surgery*. John Wiley & Sons. Retrieved from: https://books.google.co.uk/books?id=TFhSCAAAQBAJ&printsec=frontcover&dq=le+fort+1+book&hl=en&sa=X&ved=0ahUKEwipxouNzv3jAhXMfMAKHQ-BjoQ6AEITTAG#v=onepage&q=le%20fort&f=false

Halverson, A. L., Casey, J. T., Andersson, J., Anderson, K., Park, C., Rademaker, A. W., & Moorman, D. (2011). Communication failure in the operating room. *Surgery*, *149*(3), 305-310.

Hammond, K. R. (2010). Intuition, no!… quasirationality, yes!. *Psychological inquiry*, *21*(4), 327-337.

Hancock, P. A., Vincenzi, D. A., Wise, J. A., & Mouloua, M. (2008). *Human factors in simulation and training*. CRC Press. Retrieved from: http://jaredfreeman.com/jf_pubs/Freeman_AARforDistributedSims_HFinTraining_2008.pdf

Hartmann, T., Wirth, W., Vorderer, P., Klimmt, C., Schramm, H., & Böcking, S. (2015). Spatial Presence theory: State of the art and challenges ahead. In *Immersed in Media* (pp. 115-135). Springer International Publishing. Retrieved from:

https://www.researchgate.net/profile/Tilo_Hartmann/publication/281188189_Ways_to_measure_presence_Review_and_future_directions/links/5ab21f08aca27217100048ba/Ways-to-measure-presence-Review-and-future-directions.pdf

Hassan, L., Jylhä, H., Sjöblom, M., & Hamari, J. (2020). Flow in VR: A Study on the Relationships Between Preconditions, Experience and Continued Use.

Healthcareconferenceuk. (2019). *Conferences and Masterclasses.* Available from: https://www.healthcareconferencesuk.co.uk/conferences-masterclasses/a-practical-guide-to-human-factors-in-healthcare. Last Accessed: 2 Feb 2020.

Hederich-Martínez, C., López-Vargas, O., & Camargo-Uribe, A. (2016). Effects of the use of a flexible metacognitive scaffolding on self-regulated learning during virtual education. *International Journal of Technology Enhanced Learning*, *8*(3-4), 199-216.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner Jr, A. C., Zaklade, A. L., & Christ, R. E. (1992). Comparison of four subjective workload rating scales. *Human factors*, *34*(4), 429-439.

Hirschl, R. B. (2015). The making of a surgeon: 10,000 hours? *Journal of paediatric surgery*, *50*(5), 699-706.

Hoffman, R. R., Crandall, B., & Shadbolt, N. R. (1998) Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors*, 40(2), 254-276.

Holtaway, L. E., & Jackson, K. (2017). Facilitating Communication for Aviation Training and Maintenance. In *19th International Symposium on Aviation Psychology* (p. 276).

Hughes-Hallett, A., Mayer, E. K., Marcus, H. J., Pratt, P., Mason, S., Darzi, A. W., & Vale, J. A. (2015). Inattention blindness in surgery. *Surgical endoscopy*, 1-6.

Hud.ac.uk (2020). *University of Huddersfield- Search with Summons*. Available from: https://library.hud.ac.uk/. Last accessed: 2 Feb 2020.

Hull, L., Arora, S., Aggarwal, R., Darzi, A., Vincent, C., & Sevdalis, N. (2012). The impact of nontechnical skills on technical performance in surgery: a systematic review. *Journal of the American College of Surgeons*, *214*(2), 214-230. Retrieved from: https://www.sciencedirect.com/science/article/pii/S1072751511012282

Hutton, R. J., & Klein, G. (1999). Expert decision making. *Systems Engineering: The Journal of The International Council on Systems Engineering*, *2*(1), 32-45.

Hscic. (2016). *Hscic.gov.uk September Statistics*. Retrieved 20 June, 2016, from http://www.hscic.gov.uk/article/6202/Annual-census-shows-NHS-workforce-figures-for-September-2014.

HRA Decision tools (2020). *Do I need NHS REC approval*?. Available from: http://www.hra-decisiontools.org.uk/ethics/. Last Accessed: 2 Feb 2020.

IDC International Data Corporation (2019). *Augmented Reality and Virtual Reality Headsets Poised for Significant Growth.* Retrieved from: https://www.idc.com/getdoc.jsp?containerId=prUS44966319. Last Accessed: 2 Feb 2020.

Iscp (2015). Intercollegiate Surgical Curriculum Programme. Available from: https://www.iscp.ac.uk/NewsItem.aspx?enc=i9T4lAgmGWySsbMyyAbN6ALxkHt/boiR0/DJUh0XS7g=. Retrieved 20 June 2019.

Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport and Exercise Psychology*, *30*(5), 561-587.

Jacquemart, M. (2014). Le Fort 1 Osteotomy and Maxillary Distraction. Available from: https://www.youtube.com/watch?v=jUzCcEcfnUU https://www.youtube.com/watch?v=eAbGcYYepCo. Last Accessed: 2 Feb 2020.

Jamison, C. L. (2019). Surgical Perioperative Leadership: The Association Between Leadership Style and Team Job Satisfaction.

Jeffries, L. (2009). *Critical stylistics: The power of English*. Macmillan International Higher Education.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Harvard University Press.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, space, and environmental medicine*.

Kang, S. G., Ryu, B. J., Yang, K. S., Ko, Y. H., Cho, S., Kang, S. H., ... & Cheon, J. (2015). An effective repetitive training schedule to achieve skill proficiency using a novel robotic virtual reality simulator. *Journal of surgical education*, *72*(3), 369-376.

Key, C. E. J., Morris, A. P., & Mansfield, N. J. (2016). Situation awareness: its proficiency amongst older and younger drivers, and its usefulness for perceiving hazards. *Transportation research part F: traffic psychology and behavior*, *40*, 156-168.

Kim, K., Rosenthal, M. Z., Zielinski, D., & Brady, R. (2012, March). Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE* (pp. 143-144). IEEE.

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, *13*(2), 213-233.

Kirkbakk-Fjær, K., Hedelin, B., & Moen, Ø. L. (2016). Undergraduate nursing students' evaluation of the debriefing phase in mental health nursing simulation. *Issues in mental health nursing*, *37*(5), 360-366.Klein, G. (1976). Problem Solving Models and Laboratory Management. *Interfaces*, *6*(3), 97-103.

Klein, G. (1989). *Strategies of decision making*. Klein associates Inc yellow springs Ohio.

Klein, G., Borders, J., Newsome, E., Militello, L., & Klein, H. A. (2018). Cognitive skills training: lessons learned. *Cognition, Technology & Work*, *20*(4), 681-687.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986, September). Rapid decision making on the fire ground. In *Proceedings of the human factors society annual meeting* (Vol. 30, No. 6, pp. 576-580). Sage CA: Los Angeles, CA: Sage Publications.

Klein, G., & Crandall, B. (1996). *Recognition-Primed Decision Strategies*. Massachusetts institute of tech Cambridge.

lein, G. (2008). Naturalistic decision making. *Human factors*, *50*(3), 456-460.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE intelligent systems*, *21*(4), 70-73.

Klein, G., & Hoffman, R. R. (2008). Macrocognition, mental models, and cognitive task analysis methodology. *Naturalistic decision making and macrocognition*, 57-80.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). Decision making in action: Models and methods. Westport, CT, US: Ablex Publishing.

Klein, G. A., & Zsambok, C. E. (Eds.). (1997). *Naturalistic decision making*. L. Erlbaum Associates.

Klein, G. A. (2017). *Sources of power: How people make decisions*. MIT press.

Kleven, N. F. (2014). Virtual University hospital as an arena for medical training and health education. Norwegian University of Science and Technology Doctoral Thesis. Available from: http://www.diva-portal.org/smash/get/diva2:751080/FULLTEXT01.pdf. Last Accessed: 20 June 2019.

Kleven, N. F., Prasolova-Førland, E., Fominykh, M., Hansen, A., Rasmussen, G., Sagberg, L. M., & Lindseth, F. (2014). Virtual Operating Room for Collaborative Training of Surgical Nurses. In *CYTED-RITOS International Workshop on Groupware* (pp. 223-238). Springer International Publishing.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (Eds.). (2000). *To err is human: building a safer health system* (Vol. 6). National Academies Press.

Koo, B., Lee, H. G., Nam, Y., & Choi, S. (2015,). Immersive BCI with SSVEP in VR head-mounted display. In *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 1103-1106). IEEE.. In *Engineering in Medicine and Biology Society* (EMBC), 2015 37th Annual International Conference of the IEEE (pp. 1103-1106).

Korovin, L. N., Farrell, T. M., Hsu, C. H., White, M., & Ghaderi, I. (2018). Surgeons' expertise during critical event in laparoscopic cholecystectomy: An expert-novice comparison using protocol analysis. *The American Journal of Surgery*.

Latifi, R. (2016). Intraoperative Surgical Decision-Making: Is It Art or Is It Science or Is It Both?. In *Surgical Decision Making* (pp. 3-7). Springer, Cham.

Lavoie, P., Cossette, S., & Pepin, J. (2015). Development of a post-simulation debriefing intervention to prepare nurses and nursing students to care for deteriorating patients. *Nurse education in practice*, *15*(3), 181-191.

Lee, S. W., Park, J., ryum Kim, A., & Seong, P. H. (2012). Measuring situation awareness of operation teams in NPPs using a verbal protocol analysis. *Annals of Nuclear Energy*, *43*, 167-175.

Licht, D. M., Polzella, D. J., & Boff, K. R. (1993). *Human factors, ergonomics and human factors engineering: An analysis of definitions*. Dayton, OH: Crew System Ergonomics Information Analysis Center.

Lin, D. T., Park, J., Liebert, C. A., & Lau, J. N. (2015). Validity evidence for Surgical Improvement of Clinical Knowledge Ops: a novel gaming platform to assess surgical decision making. *The American Journal of Surgery,* 209 (1), 79-85.

Lipshitz, R., & Ben Shaul, O. (1997). Schemata and mental models in recognition-primed decision making. *Naturalistic decision making*, 293-303.

Lingard, L., Regehr, G., Orser, B., Reznick, R., Baker, G. R., Doran, D., ... & Whyte, S. (2008). Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anaesthesiologists to reduce failures in communication. *Archives of surgery,* 143(1), 12-17.

Luker, K. R., Sullivan, M. E., Peyre, S. E., Sherman, R., & Grunwald, T. (2008). The use of a cognitive task analysis–based multimedia program to teach surgical decision making in flexor tendon repair. *The American Journal of Surgery*, *195*(1), 11-15.

Maan, Z. N., Maan, I. N., Darzi, A. W., & Aggarwal, R. (2012). Systematic review of predictors of surgical performance. *British Journal of Surgery*, 99(12), 1610-1621.

MacDonald, A., & Schlesinger, J. (2018). Canary in an operating room: Integrated operating room music. *Proceedings of the Human Factors and Ergonomics Society Europe*, 79-83.

Maldonado-Macías, A., Realyvásquez, A., Hernández, J. L., & García-Alcaraz, J. (2015). Ergonomic assessment for the task of repairing computers in a manufacturing company: A case study. *Work*, *52*(2), 393-405. Retrieved from: https://www.researchgate.net/profile/Aide_Maldonado/publication/282278181_Ergonomic_assessment_for_the_task_of_repairing_computers_in_a_manufacturing_company_A_case_study/links/564e4dfb08ae4988a7a62269.pdf

Malhotra, M., Singh, A. V., & Matam, R. (2019, February). Comparative Performance Issues with H. 264 vs H. 265. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing,* (pp. 283-288). IEEE.

Markose, G., Berezowski, B. M., Chiu, G. A., & Cousin, G. C. S. (2020). Situational awareness in driving, and lessons for Surgeons. *British Journal of Oral and Maxillofacial Surgery*.

Mathieu JE, Heffner TS, Goodwin GF, et al. The influence of shared mental models on team process and performance. Journal of Applied Psychology 2000; 85: 273–83.

Mathur, A. S. (2015). Low cost virtual reality for medical training. In *2015 IEEE Virtual Reality (VR)* (pp. 345-346). IEEE.

Matthews, M. D., Pleban, R. J., Endsley, M. R., & Strater, L. D. (2000, October). Measures of infantry situation awareness for a virtual MOUT environment. In *Proceedings of the Human Performance, Situation Awareness and Automation: User Centred Design for the New Millennium Conference*.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, *22*(3), 276-282. https://www.biochemia-medica.com/en/journal/22/3/10.11613/BM.2012.031/fullArticle. Last Accessed: 22 Feb 2020.

McKnight, L. K., Stetson, P. D., Bakken, S., Curran, C., & Cimino, J. J. (2002). Perceived information needs and communication difficulties of inpatient physicians and nurses. *Journal of the American Medical Informatics Association*, *9,* 64-69.

MedicalEducationLeeds. (2020). *Events and Course Calendar*. Available from: https://www.medicaleducationleeds.com/event/the-3rd-urology-simulation-boot-camp/. Last Accessed: 2 Feb 2020.

Meister, D. (1999). *The history of human factors and ergonomics*. Mahwah, NJ: Lawrence Erlbaum Associates.

Menhadji, A., Abdelshehid, C., Osann, K., Alipanah, R., Lusch, A., Graversen, J., ... & Gerbatsch, I. (2013). Tracking and assessment of technical skills acquisition among urology residents for open, laparoscopic, and robotic skills over 4 years: is there a trend?. *Journal of endourology*, *27*(6), 783-789.

Mercier, E., Chagnon-Monarque, S., Lavigne, F., & Ayad, T. (2018). Objective assessment of technical skills in otorhinolaryngology–head and neck surgery residents: a systematic review. *Otolaryngology–Head and Neck Surgery*, *158*(1), 54-61.

Merritt, C., Daniel, M., Munzer, B. W., Nocera, M., Ross, J. C., & Santen, S. A. (2018). A cognitive apprenticeship-based faculty development intervention for emergency medicine educators. *Western Journal of Emergency Medicine*, *19*(1), 198.

Meyerbroeker, K., Morina, N., Kerkhof, G. A., & Emmelkamp, P. M. G. (2013). Virtual reality exposure therapy does not provide any additional value in agoraphobic patients: a randomized controlled trial. *Psychotherapy and psychosomatics*, *82*(3), 170-176.

Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. In Annett, J., & Stanton, N. A. (Eds.). (2000). *Task analysis*. CRC Press.

Militello, L., Lipshitz, R., & Schraagen, J. M. (2017). *Naturalistic decision making and macrocognition*. CRC Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Mishra, J., Allen, D., & Pearman, A. (2015). Information seeking, use, and decision making. *Journal of the Association for Information Science and Technology*, *66*(4), 662-673.

Mitchell, L., Flin, R., Yule, S., Mitchell, J., Coutts, K., & Youngson, G. (2011). Thinking ahead of the surgeon. An interview study to identify scrub nurses' non-technical skills. *International journal of nursing studies*, *48*(7), 818-828.

Moffat, S. D. (2009). Aging and spatial navigation: what do we know and where do we go?. *Neuropsychology review*, *19*(4), 478.

Moorthy, K., Munz, Y., Forrest, D., Pandey, V., Undre, S., Vincent, C., & Darzi, A. (2006). Surgical crisis management skills training and assessment: a stimulation-based approach to enhancing operating room performance. *Annals of surgery*, *244*(1), 139.

Moray, N. (2000). Culture, politics and ergonomics. *Ergonomics, 43,* 858-868.

Morris Jr, J. A., Carrillo, Y., Jenkins, J. M., Smith, P. W., Bledsoe, S., Pichert, J., & White, A. (2003). Surgical adverse events, risk management, and malpractice outcome: morbidity and mortality review is not enough. *Annals of surgery*, *237*(6), 844.

Moroney, W.F. and Moroney, B.W. (1999). Flight Simulation, in *Handbook of Aviation Human Factors,* Garland, D.J., Wise, J.A.., and Hopkins, V.D., Eds., Mahwah, NJ: Lawrence Erlbaum Associates.

Moulton, C. A., Regehr, G., Lingard, L., Merritt, C., & MacRae, H. (2010). Slowing down to stay out of trouble in the operating room: remaining attentive in automaticity. *Academic Medicine*, *85*(10), 1571-1577.

Moorthy, K., Munz, Y., Sarker, S. K., & Darzi, A. (2003). Objective assessment of technical skills in surgery. *British Medical Journal*, *327*(7422), 1032-1037.

Mizuno, Y., Narimatsu, H., Kodama, Y., Matsumura, T., & Kami, M. (2014). Mid-career changes in the occupation or specialty among general surgeons, from youth to middle age, have accelerated the shortage of general surgeons in Japan. *Surgery today*, *44*(4), 601-606.

Mushtaq, F., O'Driscoll, C., Smith, F. C., Wilkins, D., Kapur, N., & Lawton, R. (2018). Contributory factors in surgical incidents as delineated by a confidential reporting system. *The Annals of The Royal College of Surgeons of England*, *100*(5), 401-405.

Mylopoulos, M., & Regehr, G. (2007). Cognitive metaphors of expertise and knowledge: prospects and limitations for medical education. *Medical education*, *41*(12), 1159-1165.

Mylopoulos, M., & Regehr, G. (2011). Putting the expert together again. *Medical Education*, *45*(9), 920-926.

Nagendran, M., Gurusamy, K. S., Aggarwal, R., Loizidou, M., & Davidson, B. R. (2013). Virtual reality training for surgical trainees in laparoscopic surgery. *The Cochrane Library*.

Nakayama, T., Numao, N., Yoshida, S., Ishioka, J., Matsuoka, Y., Saito, K., ... & Kihara, K. (2016). A novel interactive educational system in the operating room–the IE system. *BMC medical education*, *16*(1), 1.

Nagy, A. G. (1999). Hierarchical decomposition of laparoscopic procedures. *Medicine Meets Virtual Reality: The Convergence of Physical & Informational Technologies: Options for a New Era in Healthcare*, *62*, 83

Nagpal, K., Vats, A., Lamb, B., Ashrafian, H., Sevdalis, N., Vincent, C., & Moorthy, K. (2010). Information transfer and communication in surgery: a systematic review. *Annals of surgery*, *252*(2), 225-239.

Naylor, J. (2004) Management. (2nd. ed.). New York: Prentice Hall.

Nhs.uk (2018). *Hospital admissions hit record high as population ages*. Available from: https://digital.nhs.uk/news-and-events/news-archive/2016-news-archive/hospital-admissions-hit-record-high-as-population-ages. Last Accessed: 2 Feb 2020.

NHSLA Litigation Authority (2016). *Annual Report and Accounts 2015/16*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/539498/NHSLA_report_2015-16_print.pdf. Last Accessed: 2 Feb 2020.

Nori, R., Piccardi, L., Migliori, M., Guidazzoli, A., Frasca, F., De Luca, D., & Giusberti, F. (2015). The virtual reality Walking Corsi Test. Computers in *Human Behavior*, 48, 72-77.

NumberGenerator.org. (2020*). Generate Random Numbers.* Available from: https://numbergenerator.org/. Last Accessed: 2 Feb 2020.

Nvivo (2020). Qualitative Data Analysis Software | NVivo. (n.d.). Retrieved 9 February 2021, from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Oculus.com (2020). *Oculus Rift*. Available from: https://www.oculus.com/rift/?locale=en_GB. Last Accessed: 2 Feb 2020.

Olmos, H., Gómez, S., Alcañiz, M., Contero, M., Andrés-Sebastiá, M. P., & Martín-Dorta, N. (2015). Combining Virtual Reality and Relaxation Techniques to Improve Attention Levels in Students from an Initial Vocational Qualification Program. In *Design for Teaching and Learning in a Networked World* (pp. 613-616). Springer International Publishing.

Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and anxiety*, *29*(2), 85-93.

Ovid Technologies (2019). *Welcome to Ovid Technologies.* Available from: https://www.ovid.com/. Last Accessed: 2 Feb 2020.

Okoli, J. O., Weller, G., & Watt, J. (2016). Information processing and intuitive decision-making on the fireground: towards a model of expert intuition. *Cognition, Technology & Work*, *18*(1), 89-103.

Olympusamerica.com (2020). *The World's Only Articulating HD 3D Video Laparoscope*. Available from http://medical.olympusamerica.com/products/endoeye-flex-3d. Last Assessed: 2 Feb 2020.

Oomens, P., Fu, V. X., Kleinrensink, G. J., & Jeekel, J. (2019). The effect of music on simulated surgical performance: a systematic review. *Surgical endoscopy*, 1-11.

Pagliuca, G., Monaghan, P., & McIntosh, R. (2008). Variable vulnerability of words to visual impairment: exploring grain-size effects in reading. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 30, No. 30).

Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games?. *Simulation & Gaming*, *50*(2), 136-159.

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *The Journal of general psychology*, *129*(3), 257-299.

Palter, V. N., & Grantcharov, T. P. (2010). Simulation in surgical education. *Cmaj*, *182*(11), 1191-1196. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917931/pdf/1821191.pdf. Last Accessed: 2 Feb 2020.

Park, J.K. (2010). Analysis of Communication Contents and the Associated Crew Performance Data Collected from Abnormal Operating Conditions for Nuclear Power Plants, *Korea Atomic Energy Research Institute*, KAERI/TR-4159/2010.

Parker, S. H., Flin, R., McKinley, A., & Yule, S. (2013). The Surgeons' Leadership Inventory (SLI): a taxonomy and rating system for surgeons' intraoperative leadership skills. *The American Journal of Surgery*, *205*(6), 745-751.

Parong, J., & Mayer, R. E. (2020). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*.

Parsons, T. D., Rizzo, A. A., Rogers, S., & York, P. (2009). Virtual reality in paediatric rehabilitation: a review. *Developmental neurorehabilitation*, *12*(4), 224-238.

Pears, M., Peebles, D., Ward, P., Pulijala, Y., Ayoub, A., Bagri, G., & Ma, M. (2017). Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?. *Naturalistic Decision Making and Uncertainty.*, 341.

Pears, M., Biyani, S., Joyce, A., Spearpoint, K., Van Cleynenbreugeld, B., Patterson, & Mushtaq, F. (2021). Capturing the Non-Technical Skills of a Technical Skills Trainer (NTS-TeST) during Simulation. *The American Journal of Surgery.* In Press.

Perry, S. J., Witt, L. A., Penney, L. M., & Atwater, L. (2010). The downside of goal-focused leadership: The role of personality in subordinate exhaustion. *Journal of Applied Psychology*, *95*(6), 1145.

Peshkepija, A. N., Basson, M. D., Davis, A. T., Ali, M., Haan, P. S., Gupta, R. N., ... & Anderson, C. I. (2017). Perioperative self-reflection among surgical residents. *The American Journal of Surgery*, *214*(3), 564-570.

Phipps, D., Meakin, G. H., Beatty, P. C., Nsoedo, C., & Parker, D. (2008). Human factors in anaesthetic practice: insights from a task analysis. *British journal of anaesthesia*, *100*(3), 333-343.

Piromchai, P., Avery, A., Laopaiboon, M., Kennedy, G., & O'Leary, S. (2015). Virtual reality training for improving the skills needed for performing surgery of the ear, nose or throat. *Cochrane Database of Systematic Reviews*, (9).

Prensky, M. (2001). Digital natives, digital immigrants. *On the horizon*, *9*(5).

Psnetwork, (2019). '*#TheatreCapChallenge: Where's the evidence?*'. Available from: https://www.psnetwork.org/theatrecapchallenge-wheres-the-evidence/. Last Accessed: Feb 3 2020.

Pugh, C. M., DaRosa, D. A., Santacaterina, S., & Clark, R. E. (2011). Faculty evaluation of simulation-based modules for assessment of intraoperative decision making. *Surgery*, *149*(4), 534-542.

Raheem, S., Ahmed, Y. E., Hussein, A. A., Johnson, A., Cavuoto, L., May, P., ... & Guru, K. A. (2018). Variability and interpretation of communication taxonomy during robot-assisted surgery: do we all speak the same language? *BJU international*, *122*(1), 99-105.

Rasmussen, J. (1990). Mental models and the control of action in complex environments. In: Ackermann D, Tauber MH, editors. Mental models and human-computer interaction 1. North-Holland: Elsevier Science Publishers. p. 41-6.

Rasmussen, J. (2000). Human factors in a dynamic information society: where are we heading?. *Ergonomics*, *43*(7), 869-879.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE transactions on systems, man, and cybernetics*, (3), 257-266.

Rasmussen, J. (1986). Information processing and human-machine interaction. *An approach to cognitive engineering*.

Rasmussen J. (1990). Mental models and the control of action in complex environments. In: Ackermann D, Tauber MH, editors. Mental models and human-computer interaction. North-Holland: Elsevier Science Publishers; p. 41-6.

Rcsed. (2020A). *Royal College of Surgeons Edinburgh Non-Technical Skills for Surgeons.* Available from: https://www.rcsed.ac.uk/professional-support-development-resources/learning-resources/non-technical-skills-for-surgeons-notss. Last Accessed: 2 Feb 2020.

Rcsed. (2020b). *Royal College of Surgeons Edinburgh Courses, Events, and Exams.* Available from: https://www.rcsed.ac.uk/courses-events-exams-search-results-page?Search=non%20technical%20skills%20for%20surgeons&SearchType=Course. Last Accessed: 2 Feb 2020.

Rcseng.ac.uk (2016). *Surgery and the NHS in numbers*. Available from: https://www.rcseng.ac.uk/media/media-background-briefings-and-statistics/surgery-and-the-nhs-in-numbers. Last Accessed: 20 June 2018.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 475-484. doi:10.1098/rstb.1990.0090. Last Accessed: 2 Feb 2020.

Reason, J. (1995). A systems approach to organizational error. *Ergonomics*, *38*(8), 1708-1721. doi:10.1080/00140139508925221.

Reason, J. (1997). Managing the risks of organizational accidents. Aldershot: Ashgate. ISBN 1840141042.

Reason, J. (2000). Human error: models and management. *British Medical Journal*, *320*(7237), 768-770.

Regenbogen, S. E., Greenberg, C. C., Studdert, D. M., Lipsitz, S. R., Zinner, M. J., & Gawande, A. A. (2007). Patterns of technical error among surgical malpractice claims: an analysis of strategies to prevent injury to surgical patients. *Annals of surgery*, 246(5), 705-711.

Repetto, C., & Riva, G. (2011). From virtual reality to inter-reality in the treatment of anxiety disorders. *Neuropsychiatry*, *1*(1), 31.

Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills—changes in the wind. *New England Journal of Medicine*, *355*(25), 2664-2669.

Robb, M., & Miller, G. (2012). Human factors engineering in oil and gas-a review of industry guidance. *Work*, *41*, 752-762.

Robson, W. (2017). Tools and techniques to improve teamwork and avoid patient harm. *Nursing Times*, *113*(1), 24-27.

Rose, J., Bearman, C., Naweed, A., & Dorrian, J. (2019). Proceed with caution: using verbal protocol analysis to measure situation awareness. *Ergonomics*, *62*(1), 115-127.

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: does crew resource management (CRM) training work?. *Human factors*, *43*(4), 641-674.

Salvendy, G. (Ed.). (2012). *Handbook of human factors and ergonomics*. John Wiley & Sons. Hoboken, NJ: John Wiley & Sons.

Sampene, K. C., Littleton, E. B., Kanter, S. L., & Sutkin, G. (2019). Preventing Error in the Operating Room: Five Teaching Strategies for High Stakes Learning. *Journal of Surgical Research*, *236*, 12-21.

Santander Mobility Fund Award (2019). *Santander Mobility Funds*. Available From: https://www.santander.co.uk/about-santander/media-centre/press-releases/santander-group-awards-over-36000-scholarships-mobility. Last Accessed: 2 Feb 2020.

Sararit, N., Haddawy, P., & Suebnukarn, S. (2018, July). Effectiveness of a low-cost VR simulator for emergency management training in dental surgery. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-6). IEEE.

Sarker, S. K., Chang, A., Albrani, T., & Vincent, C. (2008). Constructing hierarchical task analysis in surgery. *Surgical endoscopy*, 22(1), 107-111.

Sarker, S. K., Rehman, S., Ladwa, M., Chang, A., & Vincent, C. (2009). A decision-making learning and assessment tool in laparoscopic cholecystectomy. *Surgical endoscopy*, *23*(1), 197-203.

Savage, R. & Habinek. (1982). A multi-level Menu-Driven User Interface: Design and Evaluation through Simulation In Thomas, C. & Schneider, M. (1984). *Human Factors in Computer Systems* (pp165-186). Norwood, New Jersey: Ablex Publishing.

Scafà, M., Serrani, E. B., Papetti, A., Brunzini, A., & Germani, M. (2019, July). Assessment of Students' Cognitive Conditions in Medical Simulation Training: A Review Study. In *International Conference on Applied Human Factors and Ergonomics* (pp. 224-233). Springer, Cham.

Schaafstal, A., Schraagen, J. M., & Van Berl, M. (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(1), 75-86.

Schatzschneider, C., Bruder, G., & Steinicke, F. (2016). Who turned the clock? Effects of manipulated zeitgebers, cognitive load and immersion on time estimation. *IEEE transactions on visualization and computer graphics*, *22*(4), 1387-1395.

Schaverien, M. V. (2010). Development of expertise in surgical training. Journal of surgical education, 67(1), 37-43.

Schlickum, M., Hedman, L., Enochsson, L., Henningsohn, L., Kjellin, A., & Felländer-Tsai, L. (2011). Surgical simulation tasks challenge visual working memory and visual-spatial ability differently. *World journal of surgery*, *35*(4), 710-715. doi: 10.1007/s00268-011-0981-2.

Schober, M. F. (1998). How addressees affect spatial perspective choice in dialogue. In P. Olivier & K.-P. Gapp (Eds.), *Representation and processing of spatial expressions* (p. 231–245). Lawrence Erlbaum Associates Publishers.

Scholar (2019) Google Scholar. Available from: https://scholar.google.com/. Last Accessed: 2 Feb 2020.

Servais, E. L., LaMorte, W. W., Agarwal, S., Moschetti, W., Mallipattu, S. K., & Moulton, S. L. (2006). Teaching surgical decision-making: an interactive, web-based approach. *Journal of surgical research*, *134*(1), 102-106.

Sevdalis, N., Jacklin, R., and Vincent. (2012). Surgical Decision-Making: A Multimodal Approach. In *Safer Surgery: Analysing Behaviour in the Operating Theatre*, Rhona Flin and Lucy Mitchell (eds.). Ashgate, 353–340.

Sevdalis, N., & McCulloch, P. (2006). Teaching evidence-based decision-making. *Surgical Clinics*, *86*(1), 59-70.

Seymour, N. E. (2008). VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World journal of surgery*, *32*(2), 182-188.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, *236*(4), 458.

Shaw, L., Wünsche, B., Lutteroth, C., Marks, S., Buckley, J., & Corballis, P. (2015). Development and evaluation of an exercycle game using immersive technologies. In *Proceedings of the 8th Australasian Workshop on Health Informatics and Knowledge Management* (Vol. 164).

Sheehan, D., & Higgs, J. (2013). Practice-based education: theoretical underpinnings. In *Realising exemplary practice-based education* (pp. 13-24). Brill Sense.

Shen, Z., Tan, S., & Siau, K. (2017). Using cognitive maps of mental models to evaluate learning challenges: a case study.

Shepherd, A., & Stammers, R. B. (2005). Task analysis. *Evaluation of human work*, *3*. In Wilson, J.R. & Cortlett, N (Ed). Evaluation of Human Work 3rd ed. Florida: Taylor & Francis Group.

Shiralkar, U. (2011). *Smart Surgeons; Sharp Decisions: Cognitive skills to avoid errors & achieve results*. TFM Publishing Limited.

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: a translational study in virtual reality. *Frontiers in psychology*, *6*.

Siu, J., Maran, N., & Paterson-Brown, S. (2016). Observation of behavioural markers of non-technical skills in the operating room and their relationship to intra-operative incidents. *The surgeon*, *14*(3), 119-128.

Smink, D. S., Peyre, S. E., Soybel, D. I., Tavakkolizadeh, A., Vernon, A. H., & Anastakis, D. J. (2012). Utilization of a cognitive task analysis for laparoscopic appendectomy to identify differentiated intraoperative teaching objectives. *The American Journal of Surgery*, *203*(4), 540-545.

Spencer. F, (1978). Teaching and measuring surgical techniques: the technical evaluation of competence. *Bulletin of the American College of Surgery* 63, 3: 9–12.

Spiro, R. J. (2012). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In *Cognition, education, and multimedia* (pp. 177-220). Routledge.

Stefanidis, D., Wang, F., Korndorffer, J. R., Dunne, J. B., & Scott, D. J. (2010). Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. *Surgical endoscopy*, *24*(2), 377-382.

Stevens, S. S. (1946). On the theory of scales of measurement. American Association for the Advancement of Science.p768-770.

Sullivan, M. E., Brown, C. V., Peyre, S. E., Salim, A., Martin, M., Towfigh, S., & Grunwald, T. (2007). The use of cognitive task analysis to improve the learning of percutaneous tracheostomy placement. *The American journal of surgery*, *193*(1), 96-99.

Sullivan, M. E., Yates, K. A., Inaba, K., Lam, L., & Clark, R. E. (2014). The use of cognitive task analysis to reveal the instructional limitations of experts in the teaching of procedural skills. *Academic Medicine*, *89*(5), 811-816.

Sundar, S. S., Go, E., Kim, H. S., & Zhang, B. (2015). Communicating art, virtually! Psychological effects of technological affordances in a virtual museum. *International Journal of Human-Computer Interaction*, *31*(6), 385-401.

Surgicalsciences (2017). Oxford University surgical lectures: Non-technical skills for surgeons. Available from: https://www.youtube.com/user/surgicalsciences

Surgeons.org (2019). *Professional Development Activities- Non-Technical Skills.* Available from: https://www.surgeons.org/education/professional-development/all-professional-development-activities/non-technical-skills-for-surgeons-notss. Last Accessed: 2 Feb 2020.

Suss, J., & Ward, P. (2015). Predicting the future in perceptual-motor domains: perceptual anticipation, option generation and expertise. *Cambridge University Press.*

Suss, J., & Ward, P. (2018). Revealing perceptual–cognitive expertise in law enforcement: an iterative approach using verbal-report, temporal-occlusion, and option-generation methods. *Cognition, Technology & Work*, *20*(4), 585-596.
Stanton, N. A., Baber, C., Jenkins, D. P., Walker, G. H., Rafferty, L. A., & Salmon, P. M. (2013). *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing, Ltd.

Stanton, N. A., Hedge, A., Brookhuis, K., Salas, E., & Hendrick, H. W. (Eds.). (2004). *Handbook of human factors and ergonomics methods*. CRC press.

Stelfox, H. T., Palmisani, S., Scurlock, C., Orav, E. J., & Bates, D. W. (2006). The “To Err is Human” report and the patient safety literature. *Quality and Safety in Health Care*, *15*(3), 174-178.

Sugden, C., Athanasiou, T., & Darzi, A. (2012, September). What are the effects of sleep deprivation and fatigue in surgical practice?. In *Seminars in thoracic and cardiovascular surgery* (Vol. 24, No. 3, pp. 166-175). WB Saunders.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

Tang, B., Hanna, G. B., & Cuschieri, A. (2005). Analysis of errors enacted by surgical trainees during skills training courses. *Surgery*, *138*(1), 14-20.

Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory & Psychology*, *4*(4), 525-547.

Terema. (2020). *Team Resource Management.* Available from: http://www.terema.co.uk/. Last Accessed: 2 Feb 2020.

Thijssen, A. S., & Schijven, M. P. (2010). Contemporary virtual reality laparoscopy simulators: quicksand or solid grounds for assessing surgical trainees? *The American Journal of Surgery*, *199*(4), 529-541.

Torkington, J., Smith, S. G. T., Rees, B. I., & Darzi, A. (2001A). Skill transfer from virtual reality to a real laparoscopic task. *Surgical endoscopy*, *15*(10), 1076-1079.

Torkington, J., Smith, S. G. T., Rees, B., & Darzi, A. (2001B). The role of the basic surgical skills course in the acquisition and retention of laparoscopic skill. *Surgical Endoscopy*, *15*(10), 1071-1075.

Trickett, S. B., & Trafton, J. G. (2007). “What if…”: The use of conceptual simulations in scientific reasoning. *Cognitive Science*, *31*(5), 843-875.

Trickett, S. B., & Trafton, J. G. (2009). A primer on verbal protocol analysis. *The PSI handbook of virtual environments for training and education*, *1*, 332-346.

Tsuburaya, A., Soma, T., Yoshikawa, T., Cho, H., Miki, T., Uramatsu, M., ... & Yule, S. (2016). Introduction of the non-technical skills for surgeons (NOTSS) system in a Japanese Cancer centre. *Surgery today*, *46*(12), 1451-1455.

Unity3d.com. (2020). *Unity Personal Edition*. Available from: https://unity3d.com/get-unity. Last Accessed: 2 Feb 2020.

Unreal Engine. (2020). Unreal Engine Download. Available from: https://www.unrealengine.com/id/logout?redirectUrl=https%3A%2F%2Fwww.unrealengine.com%3FsessionInvalidated%3Dtrue. Last Accessed: 2 Feb 2020.

Van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team learning: building shared mental models. *Instructional Science*, *39*(3), 283-301.

Vannaprathip, N., Haddawy, P., Suebnukarn, S., Sangsartra, P., Sasikhant, N., & Sangutai, S. (2016, March). Desitra: A Simulator for Teaching Situated Decision Making in Dental Surgery. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (pp. 397-401).

Vassiliou, M. C., Feldman, L. S., Andrew, C. G., Bergman, S., Leffondré, K., Stanbridge, D., & Fried, G. M. (2005). A global assessment tool for evaluation of intraoperative laparoscopic skills. *The American journal of surgery*, *190*(1), 107-113.

Velmahos, G. C., Toutouzas, K. G., Sillin, L. F., Chan, L., Clark, R. E., Theodorou, D., & Maupin, F. (2004). Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory. *The American Journal of* Surgery, 187(1), 114-119

Vincent, C., Neale, G., & Woloshynowych, M. (2001A). Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal*, *322*(7285), 517-519.

Vincent, C., Neale, G., & Woloshynowych, M. (2001B). Exploring the causes of adverse events in NHS hospital practice. *Journal of the Royal Society of Medicine*, *94*(7), 322-330.

Wahr, J. A., Prager, R. L., Abernathy 3rd, J. H., Martinez, E. A., Salas, E., Seifert, P. C., ... & Sanchez, J. A. (2013). Patient safety in the cardiac operating room: human factors and teamwork: a scientific statement from the American Heart Association. *Circulation*.

Walker, G. H., Stanton, N. A., & Young, M. S. (2001). An on-road investigation of vehicle feedback and its role in driver cognition: Implications for cognitive ergonomics. *International Journal of Cognitive Ergonomics*, *5*(4), 421-444. https://doi.org/10.1207/S15327566IJCE0504_4

Wanzel, K. R., Hamstra, S. J., Caminiti, M. F., Anastakis, D. J., Grober, E. D., & Reznick, R. K. (2003). Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery*, *134*(5), 750-757.

Wanzel, K. R., Anastakis, D. J., McAndrews, M. P., Grober, E. D., Sidhu, R. S., Taylor, K., ... & Hamstra, S. J. (2007). Visual–spatial ability and fMRI cortical activation in surgery residents. *The American Journal of Surgery*, *193*(4), 507-510.

Waterworth, J., Waterworth, E., Riva, G., & Mantovani, F. (2015). Presence: Form, Content and Consciousness. In*: Immersed in media: telepresence theory, measurement & technology,* Lombard,M., Biocca, F., Freeman, J., IJsselsteijn, W., Schaevitz, R. New York: Springer-Verlag, 2015, p35-58.

Webb, R. K., Russell, W. J., Klepper, I., & Runciman, W. B. (1993). The Australian Incident Monitoring Study. Equipment failure: an analysis of 2000 incident reports. *Anaesthesia and intensive care*, *21*(5), 673-677.

Weber, J., Catchpole, K., Becker, A. J., Schlenker, B., & Weigl, M. (2018). Effects of flow disruptions on mental workload and surgical performance in robotic-assisted surgery. *World journal of surgery*, *42*(11), 3599-3607. Available from: https://link.springer.com/content/pdf/10.1007%2Fs00268-018-4689-4.pdf Last Accessed: 2 Feb 2020.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in psychology*, *10*, 158.

Westmidlandsdeanery. (2019). *Human Factors Events.* Available from: https://www.westmidlandsdeanery.nhs.uk/search-results?search=human%20factors. Last Accessed: 2 Feb 2020.

Whelehan, D. F., Connelly, T. M., & Ridgway, P. F. COVID-19 and Surgery: A thematic analysis of unintended consequences on performance, practice and surgical training. *The Surgeon*.

Who.com. (2020). *Patient Safety: Education & Training*. Available from: http://www.who.int/patientsafety/education/en/. Last Accessed: 2 Feb 2020.

Wickens, C.D, and Dehais, F., (2019). Expertise in Aviation. In: The Oxford Handbook of Expertise. Oxford University Press. ISBN 9780198795872

Wickens, C.D., & Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd ed). Upper Saddle River, NJ: Prentice Hall.

Wiegmann, D. A., & Shappell, S. A. (2017). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Routledge.Wilson, J.R. (2000). Fundamentals of ergonomics in theory and practice. *Applied Ergonomics, 31*, 557-567.

Wierwille, W. W., & Casali, J. G. (1983, October). A validated rating scale for global mental workload measurement applications. In *Proceedings of the human factors society annual meeting* (Vol. 27, No. 2, pp. 129-133). Sage CA: Los Angeles, CA: Sage Publications.

Yates, K., Sullivan, M., & Clark, R. (2012). Integrated studies on the use of cognitive task analysis to capture surgical expertise for central venous catheter placement and open cricothyrotomy. The American Journal of Surgery, 203(1), 76-80.

Yiasemidou, M., Glassman, D., Tomlinson, J., Song, D., & Gough, M. (2016). Perceptions about the present and future of surgical simulation: a national study of mixed qualitative, quantitative methodology: O55. *British Journal of Surgery*, *103*, 20-21.

Yiasemidou, M., Roberts, D., Glassman, D., Tomlinson, J., Biyani, S., & Miskovic, D. (2017). A multispecialty evaluation of Thiel cadavers for surgical training. *World Journal of Surgery*, *41*(5), 1201-1207.

Yorksandhumberdeanery (2016). *School of Surgery Training Facilities.* Available from: http://www.yorksandhumberdeanery.nhs.uk/surgery/training_facilities/. Last Accessed: 2 Feb 2020.

Young, R. S., & O'Regan, D. J. (2010). Cardiac surgical theatre traffic: time for traffic calming measures? *Interactive cardiovascular and thoracic surgery*, *10*(4), 526-529.

YouTube.com (2019). Copyright four factors information. Available from: https://www.youtube.com/about/copyright/fair-use/#yt-copyright-four-factors. Last Accessed: 3 Feb 2020.

Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., & Paterson-Brown, S. (2008). Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World journal of surgery*, *32*(4), 548-556.

Yule, S., Flin, R., Maran, N., Youngson, G., Mitchell, A., Rowley, D., & Paterson-Brown, S. (2008B). Debriefing surgeons on non-technical skills (NOTSS). *Cognition, Technology & Work*, *10*(4), 265-274.

Yule, S., Flin, R., Paterson-Brown, S., & Maran, N. (2006). Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery*, *139*(2), 140-149.

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., & Rowley, D. (2006). Development of a rating system for surgeons' non-technical skills. *Medical education*, *40*(11), 1098-1104.

Zheng, B., Swanström, L. L., & Mackenzie, C. L. (2007). A laboratory study on anticipatory movement in laparoscopic surgery: a behavioural indicator for team collaboration. *Surgical endoscopy*, *21*(6), 935-940.

Zhou, Y., Wijewickrema, S., Ioannou, I., Bailey, J., Kennedy, G., Nestel, D., & O'Leary, S. (2018). Do experts practice what they profess? *PloS one*, *13*(1).

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... & Niu, P. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*.

13 Appendix

Contents

Figure 1: Ethical Approval Form

THE UNIVERSITY OF HUDDERSFIELD

School of Human and Health Sciences – School Research Ethics Panel

Kirsty Thomson SREP Administrator: hhs_srep@hud.ac.uk

Name of applicant: Matthew Pears

Title of study: Can Expert-level Cognition be Rapidly Acquired? The Effect of a Human Factors-based Virtual Reality Trainer on Non-Technical Skills in the Operating Theatre

Department: Psychology   Date sent: 01/09/2016

Please provide sufficient detail below for SREP to assess the ethical conduct of your research. You should consult the guidance on filling out this form and applying to SREP at http://www.hud.ac.uk/hhs/research/srep/.

Researcher(s) details	Matthew Pears, 2 Blackburn Croft, Sheffield, S35 2ZJ, 07575151512
Supervisor(s) details	Dr David Peebles, Reader in Cognitive Science, AC2E research group, Ramsden Building R2, Queensgate Huddersfield, HD1 3DH, 01484 473620
All documentation has been read by supervisor (where applicable)	YES / NO / NOT APPLICABLE This proposal will not be considered unless the supervisor has submitted a report confirming that (s)he has read all documents and supports their submission to SREP
Aim / objectives	1.Significantly increase the technical and non-technical skills of surgical trainees for a maxillofacial operation, via Human Factor analysis and modification of a novel HMD surgical training simulator. In other words: 1A. To uncover cognitive differences between experts and trainees in a surgical procedure (le fort 1). 1B. To form a method to incorporate these captured expert skills into the novel surgical training simulator, for the trainees to more rapidly acquire.

Brief overview of research methods	Cognitive interviews and relevant questionnaires.
Project start date	October 2015
Project completion date	October 2018
Permissions for study	A potential 12 surgeons will be emailed and an outline of my study will be given to them. A supporting professor will contact the relevant surgeons with this information. If a surgeon shows interest, I will be given their contact details to then provide more details about the study. Their permission will be recorded via consent form.
Access to participants	Participants will be aged 18+. In collaboration with Huddersfield University PhD student Yeshwanth Pulijala, participants are recruited with help from an Oral and Maxillofacial Surgeon and lecturer from the University of Glasgow. Participants will be emailed to arrange meeting times if interest is shown. From there, meetings will be arranged to collect data in an agreed location. Participant age, sex, and data from questionnaires/interviews will be gathered.
Confidentiality	Participants' information will be confidential and a limited number of individuals will be able access data (relevant markers, and an external examiner).
Anonymity	Identifying (names, places) or signifying (history, hobbies) words will been extracted from the data to increase anonymity of participants. Anonymity of participants will be implemented throughout the entire length of the study. Finished work will have complete anonymity for all participants. Published and/or unpublished work will also anonymise participants.
Right to withdraw	Participants can withdraw at any time without consequence or reason. However, after the dissemination of the results their anonymised data will not be able to be removed from the study.

Data Storage	The raw data (questionnaire scores, interview audio recordings) will be digitally saved and encrypted and kept on a password protected computer. Physical data will be stored in a locked storage unit with access by the research only. The codes/passwords kept on another password secure computer away from the raw data. Data will be destroyed 5 years after the completion of the researcher's course at Huddersfield University, in accordance with their regulations. Data will be destroyed by shredding of all physical material, and permanent deletion of all saved electronic data.
Psychological support for participants	During their involvement, participants can communicate with the researcher at any time if they need to, via email, meetings, or telephone. From the start of contact, participants will have an information sheet to refer to. After the experiments, participants can openly ask questions and will be able to contact the researcher at any time. A debriefing form with names and contact details for relevant support will be given to participants. This includes the contact details for my supervisor and myself.
Researcher safety / support (attach completed University Risk Analysis and Management form)	Researcher safety can be maintained by informing the supervisor of any intended meetings with participants. Telephone contact can be made between the researcher and the supervisor if any issues arise. Researcher support is by email, telephone and one-to-one meetings agreed by the researcher and supervisor.
Information sheet	See figure 3
Consent form	See figure 2
Letters / posters / flyers	--
Questionnaire / Interview guide	-
Debrief (if appropriate)	See figure 4

Dissemination of results	Results of the study will be available to all participants. Participants will be asked if they would like a copy of the final results. The results may be available at Huddersfield Library for others to view. There is a possibility of publication.
Identify any potential conflicts of interest	There are no potential conflicts of interest.
Does the research involve accessing data or visiting websites that could constitute a legal and/or reputational risk to yourself or the University if misconstrued? Please state Yes/No If Yes, please explain how you will minimise this risk	No
The next four questions relate to Security Sensitive Information – please read the following guidance before completing these questions: http://www.universitiesuk.ac.uk/highereducation/Documents/2012/OversightOfSecuritySensitiveResearchMaterial.pdf	
Is the research commissioned by, or on behalf of the military or the intelligence services? Please state Yes/No If Yes, please outline the requirements from the funding body regarding the collection and storage of Security Sensitive Data	No

Is the research commissioned under an EU security call? Please state Yes/No If Yes, please outline the requirements from the funding body regarding the collection and storage of Security Sensitive Data	No
Does the research involve the acquisition of security clearances? Please state Yes/No If Yes, please outline how your data collection and storages complies with the requirements of these clearances	No
Does the research concern terrorist or extreme groups? Please state Yes/No If Yes, please complete a Security Sensitive Information Declaration Form	No

Does the research involve covert information gathering or active deception? Please state Yes/No	No
Does the research involve children under 16 or participants who may be unable to give fully informed consent? Please state Yes/No	No
Does the research involve prisoners or others in custodial care (e.g. young offenders)? Please state Yes/No	No
Does the research involve significantly increased danger of physical or psychological harm for the researcher(s) and/or the subject(s), either from the research process or from the publication of findings? Please state Yes/No	No

Does the research involve risk of unplanned disclosure of information you would be obliged to act on? Please state Yes/No	No
Other issues	Non
Where application is to be made to NHS Research Ethics Committee / External Agencies	Not Applicable
Please supply copies of all relevant supporting documentation electronically. If this is not available electronically, please provide explanation and supply hard copy	

All documentation must be submitted to the SREP administrator. All proposals will be reviewed by two members of SREP.

If you have any queries relating to the completion of this form or any other queries relating to SREP's consideration of this proposal, please contact the SREP administrator (Kirsty Thomson) in the first instance – hhs_srep@hud.ac.uk

Figure 2: Safety Assessment Form

UNIVERSITY OF HUDDERSFIELD - GENERAL HEALTH AND SAFETY RISK ASSESSMENT FORM

(To be completed for intended and proposed activities)

Brief description of activity: PC usage		
Location: Huddersfield University	**Assessment by:**	**Assessment date:**

SPECIFIC TASK/ASPECT OF ACTIVITY: Operation of display screen equipment						
Hazards identified	**Risks to health and safety**	**People at risk**	**Measures to manage the risks effectively**	**Action by:**		
				Who	**When**	**Completed**
Computer and VDU operation	Working for prolonged periods without change of posture or sufficient break.	Staff and students working at display screen equipment.	Ensure staff and student awareness of the health and safety aspects of working at display screen equipment through training and/or information.			
	Inappropriate layout, lack of awareness etc resulting in poor posture being adopted when using display screen equipment.		The workstation and equipment is subject to the risk assessment process given within the University's specific policy for ensuring health and safety whilst working at display screen equipment.			

RISK ASSESSMENT REVIEW:

To be carried out by:	Date when to be carried out by: 14/09/2016

UNIVERSITY OF HUDDERSFIELD - GENERAL HEALTH AND SAFETY RISK ASSESSMENT FORM

SPECIFIC TASK/ASPECT OF ACTIVITY: Data transport and storage						
Hazards identified	**Risks to health and safety**	**People at risk**	**Measures to manage the risks effectively**	**Action by:**		
				Who	**When**	**Completed**
Loss of Data		Participants	Ensure data is stored in secure physical locations and during travel, always being password encrypted.			

SPECIFIC TASK/ASPECT OF ACTIVITY: Interviewing participants						
Hazards identified	**Risks to health and safety**	**People at risk**	**Measures to manage the risks effectively**	**Action by:**		
				Who	**When**	**Completed**
Travel (if necessary)	Travel related	Researcher	Ensure meeting points are safe and maintain at least 1 form of communication between researcher and supervisor. Infor supervisor of planned interview times/locations			

RISK ASSESSMENT REVIEW:

To be carried out by:	Date when to be carried out by: 14/09/2016

Figure 3: Participants Information Sheet

University of Huddersfield

Participant information sheet

This sheet is for you to keep throughout your participation. If you are unclear about the study, your role, ethics, anonymity, confidentiality etc. then please refer to this sheet.

Expert Cognition Rapidly Acquired by OR Trainees- Synergistic Effects from a Novel HMD INteractive Virtual Environment Surgical Trainer

You are being asked to take part in a research study run by myself, Matthew Pears, supervised by Dr David Peebles. This study will investigate your cognition relating to a Le Fort 1 Osteotomy. This will help to create new non-technical skill training platform for trainee surgeons. In collaboration with PhD researcher Yeshwanth Pulijala, this study aims to create HMD training material to teach the different types of skills necessary to perform the procedure.

WHAT WILL HAPPEN

In this study, you will be interview by the researcher into your knowledge and experiences of a Le Fort 1 Osteotomy. This interview will orient around your cognitions when you perform the procedure. This is not an evaluation or assessment of your abilities, and will not be passed on to external bodies or third parties. The exploratory interview simply investigates your thoughts/reflections/experiences of the Le Fort 1 Osteotomy. The time commitment for this study typically is 1 hour – this can be spread between 1-3 sessions.

PARTICIPANTS' RIGHTS

You may decide to stop being a part of the research study at any time without explanation. You have the right to ask that any data you have supplied up to that point be withdrawn and destroyed (unless data has been disseminated into the final results). There will be no penalty for withdrawing from the experiment. You have the right to omit or refuse to answer/respond to any question that is asked to you, without penalty. You have the right to have your questions answered, in relation to the procedures. If you have any questions as a result of reading this information sheet, please ask the researcher before the study begins. All data will be anonymous in the final project submission.
Your participation in this study is voluntary. You will not receive any monetary compensation for your participation.

BENEFITS AND RISKS

There are no known benefits for you in this study; however, you may have a greater interest in how a Le Fort 1 can be taught via Head Mounted Display technology as a supplementary tool. There are no known risks in this study. It is not possible to provide feedback of individual scores to participants.

CONFIDENTIALITY/ANONYMITY

The data collected will contain identifying information about you. These data are your full name, age, sex, and information on your surgical knowledge/abilities. However, if there are any identifying (names, places) or signifying (history, hobbies) information, these will be extracted from the data to provide anonymity. Additionally, data is backed up in a secure, password protected and encrypted location. Data will be destroyed 5 years after the completion of the researcher's course at Huddersfield University, in accordance with the University's regulations. Data will be destroyed by shredding of all physical material, and permanent deletion of all saved electronic data.

FOR FURTHER INFORMATION

Dr David Peebles may be able to answer your questions about this study at any time. You may contact him via: D.Peebles @hud.ac.uk. Tel: 01484 473620. I am also available – Tel: 07575151512. If you are interested in the final results of this study, please contact me at matthew.pears@hud.ac.uk.

Figure 4: Completed CAPs Consent Forms

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ☑

I consent to taking part in it ☑

I understand that I have the right to withdraw from the research at any time ☑

without giving any reason

I give permission for my words to be quoted (by use of pseudonym) ☑

I understand that the information collected will be kept in secure conditions ☑

for a period of five years at the University of Huddersfield

I understand that no person other than the researcher/s and facilitator/s will ☑

have access to the information provided.

I understand that my identity will be protected using pseudonym ☑

and that I will not be identified in any report.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant: Dr. SIVA BHARANI	Signature of Researcher: M. Pears
Print: Siva.	Print:
Date: 13.01.2017.	Date: 13/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ☑

I consent to taking part in it ☑

I understand that I have the right to withdraw from the research at any time ☑

without giving any reason

I give permission for my words to be quoted (by use of pseudonym) ☑

I understand that the information collected will be kept in secure conditions ☑

for a period of five years at the University of Huddersfield

I understand that no person other than the researcher/s and facilitator/s will ☑

have access to the information provided.

I understand that my identity will be protected using pseudonym ☑

and that I will not be identified in any report.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher:
[signature]	M. Pears
Print:	Print:
Date: 9/1/17	Date: 9/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form

Student: Matthew Pears Course: Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees — Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	☑
I consent to taking part in it	☑
I understand that I have the right to withdraw from the research at any time without giving any reason * Consents to be Identified	☑ final report
I give permission for my words to be quoted (by use of pseudonym)	☑
I understand that the information collected will be kept in secure conditions for a period of five years at the University of Huddersfield	☑

I understand that no person other than the researcher's and facilitator/s will have access to the information provided.

I understand that my identity will be protected using pseudonym and that I will not be identified in any report.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher:
Dhupar	Matt Pears
Print: DR VIKAS DHUPAR	print: M. Pears
Date: 14/1/2017	Date: 4/1/2017.

(one copy to be retained by Participant / one copy to be retained by Researcher)

Figure 5: Completed CAPs Consent Forms (Participants not used as they were not CAPs)

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	☑
I consent to taking part in it	☑
I understand that I have the right to withdraw from the research at any time without giving any reason	☑
I give permission for my words to be quoted (by use of pseudonym)	☑
I understand that the information collected will be kept in secure conditions for a period of five years at the University of Huddersfield	☑
I understand that no person other than the researcher/s and facilitator/s will have access to the information provided.	☑
I understand that my identity will be protected using pseudonym and that I will not be identified in any report.	☑

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher:
Bamadive	M. Pears
Print:	Print:
Date: 23/1/17	Date: 23/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	☑
I consent to taking part in it	☑
I understand that I have the right to withdraw from the research at any time without giving any reason	☑
I give permission for my words to be quoted (by use of pseudonym)	☑
I understand that the information collected will be kept in secure conditions for a period of five years at the University of Huddersfield	☐
I understand that no person other than the researcher/s and facilitator/s will have access to the information provided.	☑
I understand that my identity will be protected using pseudonym and that I will not be identified in any report.	☑

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher: M. Pears
Print:	Print:
Date: 23/1/17	Date: 23/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ☑

I consent to taking part in it ☑

I understand that I have the right to withdraw from the research at any time ☑

without giving any reason

I give permission for my words to be quoted (by use of pseudonym) ☑

I understand that the information collected will be kept in secure conditions ☑

for a period of five years at the University of Huddersfield

I understand that no person other than the researcher/s and facilitator/s will ☑

have access to the information provided.

I understand that my identity will be protected using pseudonym ☑

and that I will not be identified in any report.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher: M. Pears
Print:	Print: Matthew. Pears
Date: 18/1/17	Date: 18/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

P8
mumbai
1st expt

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	☑
I consent to taking part in it	☑
I understand that I have the right to withdraw from the research at any time	☑
without giving any reason	
I give permission for my words to be quoted (by use of pseudonym)	☑
I understand that the information collected will be kept in secure conditions	☑
for a period of five years at the University of Huddersfield	
I understand that no person other than the researcher/s and facilitator/s will	☑
have access to the information provided.	
I understand that my identity will be protected using pseudonym	☑
and that I will not be identified in any report.	

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher:
[signature]	M. Pears
Print: Dr. N. N. Andrade	Print: Matt Pears
Date: 23/1/17.	Date: 23/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form

Student: Matthew Pears **Course:** Psychology PhD (Doctor of Philosophy)

Surgeons' Situational Awareness and Decision-Making Rapidly Acquired by Trainees – Can Human Factors and Modern VR Create Synergistic and Predictive Effects?

It is important that you read, understand and sign the consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	☑
I consent to taking part in it	☑
I understand that I have the right to withdraw from the research at any time without giving any reason	☑
I give permission for my words to be quoted (by use of pseudonym)	☑
I understand that the information collected will be kept in secure conditions for a period of five years at the University of Huddersfield	☑
I understand that no person other than the researcher/s and facilitator/s will have access to the information provided.	☑
I understand that my identity will be protected using pseudonym and that I will not be identified in any report.	☑

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.

Signature of Participant:	Signature of Researcher:
[signature]	M. Pears [signature]
Print:	Print:
Date: 13/1/17	Date: 13/1/17

(one copy to be retained by Participant / one copy to be retained by Researcher)

Figure 6: NTS Questions added into the assessment questionnaire during collaborative data gathering period

10. **How would you maintain the increased alar base width to its original pre-surgical dimension?**

Cinch stitch b. Apply pressure on the mucosa c. Mucosal suturing d. VY Closure

11. **During a Le Fort I Operation something unexpected happens. There is no procedure available. You have seconds to make a decision to attempt to prevent a bad outcome, what is the best thing to do?**

a. Consider all the possible options, and select the best.
b. React the best you can to the situation based upon your experience.
c. Ask your supervisor/teacher to take over and make the decision.

12. **The surgeon is making a decision during a Le Fort I that you believe will cause a small error/issue. You are aware of an important piece of information that the surgeon may have missed that may change their decision. What should you do?**

a. Trust that the surgeon did not miss/overlook this piece of information
b. Interrupt the surgeon and inform him/her before they make the decision
c. Ask a team member if they think the piece of information is important
d. Watch to see if an error occurs and after operation explain why it may have occurred

13. **How do your senses help you when opening the smith's spreaders to down-fracture the Maxilla? What can you feel, hear, see?**
Answer:

Figure 6A: CAP Probing questionnaire

Section 1 (Approximately 5 minutes)**:**

Q1: **What is your** Age**?**

Q2: **What** Sex **are you?**

Male Female Other

Q3: **How many times have you approximately** performed **a Le Fort 1 procedure?**

a. 1-9 b. 10-29 c. 30-49 d. 50-69 e. 70+

Q4: **How many times have you** taught **all, or most, of a Le Fort 1 procedure?**

a. 1-9 b. 10-29 c. 30-60 d. 50-69 e. 70+

Q5: **List 4 or more** methods of educational delivery **starting from the best to** teach students **a Le Fort 1 Osteotomy** (e.g. Lectures, 3D Animations, Books, Interactive websites, Operating Room lessons, Videos)

1st 2nd 3rd 4th

5th 6th

Q6: **Of the above teaching methods, which 2 would you like to become more available?**

Q7: **What technology have you used to implement educational material into oral and maxillofacial education?** (E.g. Computer, Tablet, Smartphone, Projectors, Simulators, replica equipment, Virtual/Augmented Reality devices)

1. 2. 3.

4. 5. 6.

Q8: **Which statement best describes** your current practice **of technology in Oral and Maxillofacial education?**

I do not include use of a tablet, PC, or smartphone to aid in surgical education ☐

I sometimes include use of a tablet, PC, or smartphone to aid in surgical education ☐

I always include use of a tablet, PC, or smartphone to aid in surgical education ☐

I depend upon use of a tablet, PC, or smartphone to aid in surgical education ☐

Q9: **Which statement best describes** <u>your knowledge</u> **about Head Mounted Displays?**

I have never heard of Head Mounted Displays ☐

I have heard of Head Mounted Displays, but I do not know what they do ☐

I know a little about Head Mounded Displays work ☐

I know a lot about Head Mounded Displays and their uses ☐

Q10: **Which statement best describes** <u>the usage</u> **of Head Mounted Displays?**

Our facility does not use Head Mounted Displays in research/education ☐

Our facility intends to use Head Mounted Displays in research/education ☐

Our facility uses Head Mounted Displays in research/education ☐

Q11: **Which statement best describes** <u>the usage</u> **of Head Mounted Displays?**

Our facility does not use Head Mounted Displays in research/education ☐

Our facility intends to use Head Mounted Displays in research/education ☐

Our facility uses Head Mounted Displays in research/education ☐

Q12: <u>If you have</u> **used a Head Mounded Display, have you** <u>viewed medical/surgical</u> **content in it? If so, what was the content?**

Q13: <u>If you have not used</u> **a Head Mounded Display, would you be interested in finding out** <u>more information</u> **about them, to aid in research/education?**

Yes ☐ / No ☐ / I do not know enough about them ☐

Section 2 (Approximately 10 minutes)**:**

These next questions will ask you about your experiences of performing a Le Fort 1 procedure.

Q14. **In no less than 3 tasks, but no more than 6 tasks, separate and provide the** most important tasks **in a Le Fort 1 Osteotomy**. (E.g. Horizontal Maxillary Osteotomy, Ensuring Maxilla Mobility, Pterygomaxillary Dysjunction, Downfracture of Maxilla, Plate Security, Ensuring/checking Occlusion etc)

2.	2.
3.	4.
5.	6.

Q15. **Which task in a Le Fort 1 Osteotomy is the most** difficult to perform? *(Any task, however general or specific)*

Q16. **Which task in a Le Fort 1 Osteotomy is the most** difficult to teach? *(Any task, however general or specific)*

Q17. **What are the** difficulties **that come to mind that make this task** difficult to teach**?** (E.g. poor visual field, complex instruments, dependency on others to perform, minimal haptic feedback, multitasking demands, decision-making demands, difficult to verbalise what you're doing, need to 'plan ahead' etc.)

Q18. **What do you believe is a** main factor **that can** prevent trainees **from proficiently learning this task?**

Q19. **In your experience,** what is the main action **that trainees do differently to you in this task?** (For example, hold the equipment differently, too slow/fast, cannot make decision for next course of action, is not aware of other important factors/issues).

Q20**. When identifying good performers from poor performers, what behaviours may be different in this difficult task?**

Q21. **Generally, what part of a Le Fort 1 Osteotomy are** <u>human errors mistakes</u> **most likely to occur?**

Q22. **How does your expertise help you to prevent errors in this task?** (Think back to a time where your expertise helped to spot an error that a trainee may not have known about)

Figure 7: Planned Semi-Structured Interview question to use with CAPs- These questions prompted many other improvised questions, but the general flow was kept where possible.

Semi-Structured Interview – SME's (Researcher's copy)

The questions in this interview will ask about experiences of performing a Le Fort 1 Osteotomy, and will also probe into abilities in decision-making, situational awareness, communication and other skills. The structure of the interview orients around their cognitions, and although some questions may be obscure or difficult to answer, they should try to answer as in-depth as possible.

General Task Identification Questions:

(3m) 0A. In no less than 3 tasks, but no more than 6 tasks, separate and provide the <u>most important tasks</u> **in a Le Fort 1 Osteotomy**. (E.g. Horizontal Maxillary Osteotomy, Ensuring Maxilla Mobility, Pterygomaxillary Dysjunction, Downfracture of Maxilla, Plate Security, Ensuring/checking Occlusion etc.) (Aim: to directly identify key tasks in the procedure)

(1m) 0B. What task is the most <u>difficult to perform</u>? *(Any task, however general or specific) (*Aim: to identify the main task of focus -for creating training tools later)

- What knowledge should a person possess to be able to conduct this procedure? (Aim: to find out the required knowledge composition of the surgeon)
- How much time did you spend in training and on the job before you really knew it? (Aim: to understand current length of time to train, to later compare deliberate *time* as an outcome of intervention)
- What are the consequences if the task is performed incorrectly or is not performed at all? (Aim: to understand the severity of mistakes, and ensure proactive training to avoid these possibilities)

(5m)0C. What are the <u>difficulties</u> **that come to mind that make this task** <u>difficult to teach</u>**?** (E.g. poor visual field, complex instruments, dependency on others to perform, minimal haptic feedback, multitasking demands, decision-making demands, difficult to verbalise what you're doing, need to 'plan ahead' etc.) (Aim: to identify current limitations of task(s) and then facilitate contextual cognitive probing for how these issues can be solved via training)

(5m) 0D. In your experience, <u>what is the main action</u> **that trainees do differently to you in this task?** (For example, hold the equipment differently, too slow/fast, cannot make decisions/unsure for next course of action, is not aware of other important factors/issues) (Aim: identify perceived differences in terms of behavioural outcomes, to be able to target cognitive training around changing these behaviours)

(3m)0E. Summary Question: Whilst performing and teaching Le Fort 1 Osteotomies in the OR, your work involves a need to analyse and focus on cases in detail. But we know that experts are also able to maintain a sense of the 'big picture' of the events occurring around them. Can you provide me with what is important about the maintaining a 'big picture' for this task? (Aim: to let them provide a general overview of key points of the environment, this should generally match up with what they have said in previous questions)

We will now discuss the task that you have identified as difficult to perform (theoretically teaching difficulty should match. If not, continue with performance task**).**

Situational Awareness Questions (choose 1 task) (*numbers refer to the task number in the Task Analysis*)

1. **(Task 10B) Maxillary Osteotomy (from pterygomatic junction to pyriform aperture)**

1. (**SAL2 Task 10.2 2m)** How should the saw be used when you initiate osteotomy from zygomatic buttress posteriorly to the lateral corner of the pyriform aperture anteriorly
2. **(SAL2 Task 10.2.1.4.2 2m)** What hand movement do you execute when you initialize the saw cut in an anteroposterior direction?
3. **(SAL3 Task 10B 5m)** What are you mindful of, that may (un) predictably occur (e.g. the saw may jump, or the cut line may not be high enough on the maxillary curve) (most prevalent errors with active attentional deployment)
4. **(SAL3 Task 10B 5m)** Have you stopped a novice just before they have made a mistake, if so how did you know they were about to make a mistake? (Cues prior to error, any type)
5. **(SAL3 Task X 5m)** Why were/are these cues useful? What information do they give you? Why do you need to find/use/look at these cues?
6. **(SAL2 Task X 3m)** Why might a novice miss these cues?

2. **(Task 14) Downfracture Maxilla using Smith's Spreaders**

1. **(SAL1/2 Task 14.1)** Where should the smith's spreaders be placed, and why?
2. **(SAL1 Task 14.1)** What cues, information etc. are you looking for to ensure spreaders are in the right places?
3. **(SAL2 Task 14.2)** What are the consequences of placing the Smith's Spreaders too far anterior?
4. **(SAL2 Task 14.2)** How do you know the Smith's spreader has been placed on the thin bone of the zygomatic buttress? How would you know a novice has done this?
5. **(SAL3 Task 14.3)** What mistakes are likely during this point?
6. **(SAL3 Task 14.3)** Do you imagine the possible consequences of this action?
7. **(SAL3 Task 14.3)** How might/has a novice behaved differently to you when using Smith's Spreaders?
8. **How does it feel when opening the smith's spreaders to downfracture the Maxilla? What can you feel, hear, see?**

7. *Think of a time where the Smith's Spreaders usage has caused error or an unexpected error has occurred:*

1. How did you know that the opening of the smith's spreaders was going wrong? (mental model and cues to aid projection)

2. What information or cues are telling you that the opening of smith's spreaders was not going correctly? (specific cues that can later train novices to look out for)

3. **(Task 15) <u>Pterygo-maxillary Disarticulation</u>**

1. **(SAL2 Task 15.1)** How do you decide the best locations to place the forceps?

2. **(SAL3 Task 15.2)** What is the goal of performing "figure of 8" circular movements?

3. **(SAL0 Task 15.2)** What do you feel when you do it?

4. **(SAL3 Task 15.2)** What are you wanting/anticipating will occur when you are moving maxilla with Rowe's?

5. **(SAL2 Task 15.2)** How do you know to stop using Rowe's forceps? What cues tell you to finish?

6. **(SAL2 Task 15.2)** Have you seen novices do it wrong?

7. **(SAL2 Task 15.2)** What do you think they didn't understand?

8. **(SAL2 Task 15.2)** As you were training what didn't you understand but the expert did?

9. **(SAL2 Task 15.2)** How are you able to tell if the bony connections have separated?

10. **(SAL2 Task 15.2)** What does it feel/look/sound like and how can you check the connections have been separated?

11. **(SAL3)** What could go wrong and how do you prevent that?

13. **During a Le Fort I Operation something unexpected happens. There is no procedure available. You have several seconds to make a decision to attempt to prevent a bad outcome, what is the best thing to do?**

a. Consider all the possible options, and select the best.
b. React the best you can to the situation based upon your experience.
c. Ask your supervisor/teacher to take over and make the decision.

14. **The surgeon is making a decision during a Le Fort I that you believe will cause a small error/issue. You are aware of an important piece of information that the surgeon may have missed that may change their decision. What should you do?**

a. Trust that the surgeon did not miss/overlook this piece of information
b. Interrupt the surgeon and inform him/her before they make the decision
c. Ask a team member if they think the piece of information is important
d. Watch to see if an error occurs and after operation explain why it may have occurred

15. **How do your senses help you when opening the smith's spreaders to down-fracture the Maxilla? What can you feel, hear, see?**

Figure 8A: Rationale for the 3 CAPs Ranking Orders

CAP 1: (order 2,3,1 ie C is best, then A, then B)

Option B yeah I have no idea what time or day this is but if you're going around looking you could end up from 5 more trainees, or yourself a media visitor who is in scrubs, its good if it's an ENT or an Otolaryngologist but its pot luck. Sometimes you ask for a tool and people are gone 15 minutes and they have been trying to find one so it's a risk and we know the issue so let's just deal with it using the people in the room we know are good. We all have blind spots but together we help prevent each other's issues, too many people is firstly too noisy and if there are 2 or 3 opinions that will take time and by then I'm sure it will have turned worse.

CAP 2: (order 2,3,1)

I am theatre nurse in this example, but I will still think like a surgeon as we should be thinking the same. And, you know, it says the other staff are very good at their job so that makes me think I most likely will be too in the example, just speaking out loud. I don't agree with the 7-minute wait here but I understand the point and this is why option A is safer . I don't understand option B so I wont comment but Option A definitely will stop the patient from enduring time like this compared to B. What is C…erm...is the same. 1 second. {*reads C 2 times I think as he takes a while}. Ok so I revise my answer I didn't read the detail of this last option here and I imagined this would happen in Option A but now this option has you making everybody aware, I would think to do this anyway, and I would judge this to be critical. Your question says it is not but an attempt for several minutes means the ceasing is causing issue that is indefinite and I don't know if its above the Larnyx or an issue further down so lets just be critical. It has happened where nurses take over especially with younger surgeons so in this scenario this is ok to do. I can say more on Option B?.. People panic., yes, easily, its easy to get someone else to help you think instead of you doing it but in the long run this doesn't make you better. we have enough to think even just 3 people. if we are calm you learn to focus better actually like this so this, for me, is not necessary.

CAP 3 (order 1,3,2)

Option A is better than C because why would I want to change something that is not yet so bad. It all sounds bad on the surface, but it says right there 'patient status is about the become critical'. Which tells me it's not critical. Which tells me I don't need more people in the room. It does tell me to be prepared for when things do, and let's assume they do, go critical. Then I need this tool. So yes option 3 is a choice I would have picked if another 30 seconds had passed in the scenario. Even on paper, it's easy to rush into a problem, give the surgeon a chance by the time I am back he either has or has not, either way I am ready for the next part and we are all on the same page of using the tool. If you get more people, then everyone is thinking and seeing different things and more chance of confusion. I'd keep it simple. Calm. But thinking ahead just in case.

Figure 8B: CAP1 Section of interview transcript

What's the most difficult part of a le fort 1 osteotomy?

P: It's fracturing the maxilla, downfracturing the maxilla.

I: What makes that difficult?

P: The total surgery depends on its entirety because if the cuts you have taken aren't complete, you're not going to achieve a fracture of the maxilla, you won't disengage it from the cranial base, so unless it is disengaged. And that's the time where you can expect a lot of, erm if it's not done properly the fracture is going to be irregular and is not going to be in a position you want. So, all those things have to be at that particular time.

I: So, it all depends on the cut?

P: It all depends on the cuts.

I: Right, it's difficult because if the cuts aren't done right then there are a lot of problems. So how can you help to reduce those difficulties? How do you stop it becoming difficult?

P: You recheck whatever you have done, you revisit, once you have given a cut, erm, when you've given a particular cut you revisit that cut, again once you've finished all the cuts you go back to the origin and start again, and see where all these have been completed.

I: Do you revisit the cuts straight away afterwards or do you try and downfracture first

P: What I do is I normally give the cuts, come back to the cuts again and see that their complete – because some amount of mobility would start even if you have not completed your cut, erm the fracture, there would be some amount of mobility into the maxilla already, but if the cuts are not complete, my amount of mobility is not there. If I see that kind of mobility I know more or less my cuts are complete, then a can go in for, you know, putting more pressure to fracture it, otherwise I would try doing that.

I: How do you check for good/correct cuts, I have no background so I am interested to know how...

P: What happens is that there are certain blind areas where you can't see, those bling areas are more important to check. Like, it's the pterygoid plates we have. What I would generally do is put my finger on the palatal site and see the movement, the general movement of the Osteotome coming and I can feel around at least that side, without reaching. If I asked my colleague or somebody to put their, erm, you know, usually you need both the hands for the tap and mallet and what I would do is ill ask somebody to tap from their side and I put my finger in so I know I have reached the…

I: So, you're holding the chisel?

P: Chisel, and feeling for where I am exactly reaching, so if I reach that particular point, you know, because you can't perforate the tissue, soft tissue, so you have to maintain the soft tissue, so if you are reaching that particular point, I know I am fine with those cuts. You need additional hands when you are working, erm. You need an assistant; an assistant probably would put a finger and check that you have made the cuts. Initially I would make him hit, because when I am closer to it, then I would move my hands in and make him tap because I don't want excessive forces during the tap, and not get excessive force as that force can go through and those areas are prone for a high amount of bleeding.

I: Right, so if a trainee was doing it, if you let them have the chisel, and check, how would you tell them? How would you describe it to them on what they should feel for?

P: Normally at first, I'll tell them it's not like a normal hard bone. That area is not going to be like a normal hard bone. It's relatively soft and an amount of force is required as relatively, the tap are relatively much and you have to hear for the change in the sound, as you keep tapping, when the, when it's gone through the hard bone and its going into the areas where the sound is going to change and get deeper and come from a sharp sound to a much duller sound and you will get that feel, you will keep feeling with the Osteotome all of the time, its, take your time and do it rather than, you know do it and be in a hurry to give one heavy force and come out with a problem.

I: What does it feel like? It might be hard to explain but what is the feel?

P: It's the feel that, you know, when you give a final tap it's like a sinking feel that it's just gone through, it means you're not in the bone now, but you have just gone through it.

I: So, you have to slow down as it gets duller go slower?

P: Yes, slow down and go softer on your taps.

I: How many taps do you usually do?

P: It depends on the patient's bone, on the density of the bone because you cannot quantify how many taps on the bone, it depends on the density. But my idea would be smaller easier taps rather than one blow.

I: It's safer to do smaller taps?

P: Yes

I: Ok thanks. My questions get stranger so bear with me.

P: It's not a problem

I: If you couldn't feel the chisel, how would you manage without feeling it? Like if you wore a glove.

P: There has never happened where I haven't felt the chisel, I have to feel it, because I know I am through. Or if you're really not feeling it then somebody could come and try to downfracture, and see the movement maybe. Come and check your, come and feel for the fracture.

I: That's like a backup, that if you have not quite completed the dysjunction, the downfracture can tell you.

P: Yes, the downfracture can.

I: If you try and downfracture and its not quite right, how many times have you gone back to the dysjunction?

P: Very rare it is, because there are 2 modifications to le fort 1, either you go through the pterygoid plates, that's the plate behind which we really need feel and be very careful about. Or we use a modification and we extract the third molar and go through that, and it makes it relatively safer and easier for us, and the areas is susceptible for palpating and feeling rather than going behind the area much behind.

I: Right. Do trainees perhaps have to check the Maxilla more than once?

P: Yes, they will have to check more than once, definitely.

I: How long does it take to check, obviously, you have to stick to a time schedule, if somebody is learning about it…

P: I think learning about it initially, the time should be a constraint. And however long you are going to take it will not take more than a particular time because that area is small where you are working and it eventually will happen but the time is not a, I would say that time should not be a parameter to be looked at.

I: So, let's say they had to do it twice

8: 52

Figure 8C: CAP2 Interview, along with an initial assessment of researcher to find if they had expertise.

I: How many years have you been performing Le Fort 1 Osteotomies?

P: I've been performing lefort 1 for about 15 to 18 years.

I: How many le fort 1's do you perform in a week or month?

P: Altogether I do about 100 osteotomies a year, and that's a combination of le fort 1 and other surgeries together.

I: Do you teach, so do you let people assist you?

P: Yes, all the cases which I do, it is always assisted by a trainee usually, so they learn from me on all the cases on a 1 to 1 basis.

I: One trainee to you, so you just have 1 trainee

P: Erm no there will be two trainees with me, at any given time when I perform a surgery, there will be two trainees assisting me, one on the right side and one on the left side, and they get to see what I do. And I talk them through part of the surgery, a simpler cut, so that they get a, erm, one to one experience of the bone feels, and go through the steps and what kind of operation it needs to be, that kind of erm method.

I: Right. Have you heard of non-technical skills before?

P: Yes, I've heard about non-technical skills, erm how should I say, the operating room, the skills, I've heard about it, but never had any experience by myself.

I: Yeah, no-one's ever taught you?

P: No, I've just had experience on my own but I have not had an experience on the non-surgical way of learning, a virtual way of learning.

I: So, we will start off generally, in le fort 1 what is the most critical task, either critical to perform or critical in the errors, what do you think is the most critical task

P: While operating?

I: Yes.

P: While operating I feel the posterior dysjunction, that is the tuberosity cut at the pterygomaxillary dysjunction, I feel is number 1, most other trainees find it most difficult to understand and the second one they find the most difficult is the erm, posterior portion of the lateral wall of the nasal cavity as you go posterior how far posterior should one go. They don't get the judgement of that. So, these are the two things I find, that the trainees find difficult.

I: What do you find most difficult?

P: [5 seconds pause] the dysjunction.

I: Ok, and just to make sure I understand what that means, the dysjunction is done before the downfracture?

P: Yes.

I: And you would use the Chisel and the Mallet?

P: Yes, the curved pterygoid Chisel, and then I do the osteotomy in the posterior part of the maxilla, to separate the pterygoid plates from the maxilla. And then I downfracture it after that.

I: Just on a side note from our interview, some other places may not do the pterygomaxillary dysjunction, do you know why?

P: Yes, it just so happens that the thing is, erm, I do also osteotomy and lots of these patients who do have cleft palate deformity and generally I have observed that in cleft lip palette deformity cases the dysjunction is more difficult and to get the downfracture is more difficult. So, in those cases you would have to do a dysjunction. Otherwise the maxilla doesn't really come down. So, that's the reason. I do a dysjunction in almost all the cases. But yes, I think in some of these cases you can get away without dysjunction. As long as they are normal orthographic surgical cases. What I mean to say, they are non-syndromic cases.

I: Even In normal cases you can still do the dysjunction?

P: Yes, and I still do it, because I generally feel the downfracture all goes very clean. And I don't use the disimpaction forceps for dysjunction, I always use the finger pressure to do the dysjunction. When I do the posterior cut, I can do the downfracture with two finger pressure. I feel it is more convenient and more comfortable and you have precise cuts. And also, the chances of fracture going, you know, cranially and causing the, a, a fracture to other bones, is not there.

I: Ok, so you said about the pterygomaxillary dysjunction being quite a critical part, what makes it critical?

P: 2 things, generally though we talk about pterygomaxillary junction as a [sutural?] separation, but mostly we find that in others it is fused. I see than in about more than 60% of the cases. I see the dysjunction is almost fused. And 2, the orientation of the intermaxillary vessel is fairly closed. And 3rd one is, you are going to operate in this particular area almost blind when you're doing the dysjunction. Because you can't visualise where exactly you're doing the dysjunction most of it is tactile sensation because you want to keep one finger intra-orally in the hammerous region, you're placing the Chisel externally and the third person is going to, you know use the Osteotome and, you know, hammer it. So, it's quite a blind procedure.

I: Can you see the Chisel end?

P: I can place the Chisel but I cannot see the end.

I: So, no one can see it?

P: No one can see it. It's very difficult to see exactly where the Chisel is going, its mostly tactile sensation.

I: Right, how do you try and solve those problems? To reduce the change of anything going wrong

P: Erm, [3-second pause], generally we have different sizes of osteotomies, and this is what I do, I'm just telling you, traditionally, in some of the cases where I am really worried I use an oscillating saw and I put it in the dysjunction area and I run it a couple of times. So, that the area becomes slightly weakened.

I: So, you start off in the bottom,

P: In the junction, yes, just behind the tuberosity, where the tuberosity ends, I put oscillating saw here and then I just cut it so that it becomes slightly weakened. And second is, the other size Chisel is a nigger one, whereas there is a smaller size available in the pterygoid Chisel, so I use a smaller one to work my way through in the dysjunction.

I: And then the bigger one.

P: And then the bigger one so that I know that the direction that I want the Chisel to go is going in the same direction. So, these are a couple of modifications which I do.

I: Yes, they sound like shortcuts, really good and creative ways.

P: More creative ways of doing it yes. And I've also tried a few cases where the third molar tooth is removed and I can go through the socket but then, I don't know, for some reason the Chisel doesn't seem to go into the third molar socket it always goes behind the socket, I generally find it more difficult to go into the third molar socket and achieve a dysjunction. It almost always slips back and goes into the dysjunction area. Or at the dysjunction of the pterygomaxillary area. So, it tends to go there, so there is an Osteotome called a 'shark fin' Osteotome, it has got an additional projection, which, when you do the le fort 1 osteotomy cut you put the shark fin osteotomy line and then you give the angulation and tap it. So, that fin will kind of guide you. So, these are a few things I've tried, but then finally I think it is the oscillating saw and working with a smaller Osteotome and then increasing it a little bit, is the one which works the best.

I: Great tenacious, so when the people shadowing you are helping, do they hold the mallet and Chisel, or… what is the setup?

P: I hold the Chisel and it will be my finger in the [hamular] area, they do the malleting. They just tap it, because I feel more confident when I am controlling the direction and the, I know how much force they are applying so I can tell them 'no, go slow'.

I: What do you usually tell them with the Chisel? Is it quite a few taps with the Chisel or it is one big one?

P: Erm, in some cases its 3 or 4 taps, in some cases it might be 2 taps.

I: Right, are they quite big taps?

P: Erm, it's not a lot of force, its, erm, [4-second pause], its.

I: [laughs] It hard to explain isn't it

P: [laughs] yes, how to explain that.

I: Is it better to do smaller taps?

P: It's better to do a smaller tap, it's more controlled, and the tap is not a single one, it is two taps.

I: How come you do two taps?

P: Erm, [pause 5 seconds] [laughs]

I: [laughs]. [Pause 3 seconds]. What does it feel like?

P: The ones in the tap, generally what happens is. I see the trainees when they put one single tap they tend to use a lot of force. You know, so it goes a little uncontrolled. So, when I tell them to do two taps so they tend to use the force in a more duratious manner. So, that probably is the reason why I tell them to do two taps.

I: Yes, I get you, to distribute the force

P: It distributes the force a little, not slowly but in a more definite manner. One single tap is generally more muscle and they don't tend to gauge how much force is needed. So, 2 taps probably, that's how I teach it.

I: Right ok, erm,

P: One more thing I forgot to tell you about the pterygomaxillary dysjunction is that I have done some cases who are 54 and 60 years when I've done le fort 1 osteotomy. For a case of tumour removal as an osteotomy. So, for the neurosurgeon to remove a tumour that in in the skull base. So, in those patients I found that at times it is just not possible to separate at the junction. So, I tried to use the Osteotome, it wouldn't go in. so I tried about 8 to ten times I was not able to get the Osteotome to come across and I could not feel it on the palatal side. As the Osteotome advances we get the feel that the Osteotome is coming through and that's how we know the dysjunction is complete. I was not able to feel it, so what I did in those cases is, the maxillary tuberosity end here [is now using a skull as a pointer] so I put the Osteotome through the maxillary tuberosity and I fractured off the lower half of the pterygoid plate together. The cut released the maxilla and I was able to downfracture the maxilla.

P: So, age is a factor I find the dysjunction is more difficult as age advances.

I: Right, so apart from this talk of the procedure itself. What else do you need to keep a check on? What do you need to know in the OR?

P: The other thing I watch for is the descending palatine vessel. When I downfracture it, then usually you would end up with a small spur of bone in the lateral wall of the nasal cavity, and that is where you would find the descending palatine vessel. You can see than. And especially when I have to remove the bone around that area for advancing the maxilla, or for superior repositioning, I take care to make sure the vessel is not damaged.

I: How do you do that?

P: I take a spatula elevator and then just protect the vessel and then using the bone nibbler, I cut the bone around the vessel.

I: You move it?

P: Yes, it is possible to get a little [SOMETHING] between the bone and the descending palatine vessel. And then we can nibble off the bone around there. So that the vessel can be saved as much as possible of course if you have to ligate** the vessel there is no other option we coagulate the vessel, we cut it up and if you move the maxilla superiorly if you move the maxilla, quite a bit forward, say about 7 or 8 millimetres in that case then you have to sacrifice the descending palatine but wherever it is possible I try to save it.

I: Do you electro-cortorise it?

P: I just cortorise it and cut it, I just buzz it, once on top once down, and cut in-between. I don't put a ligature wire thread around that.

I: When I have been talking to others, they have said the most important part has been the downfracture, but it seems they really meant the checks just before the dysjunction. Could you further talk me through the pterygomaxillary dysjunction and also the nasal mucosa, being the checks before the downfracture?

P: When I first make the incision, initially we do the superior dissection to expose the infraorbital nerve then we go posteriorly with the periosteal elevator to identify the region of the pterygomaxillary dysjunction area. Once we do that we come anteriorly to the nasal mucosa dissection. And then when I am doing this dissection it is essential to raise the nasal floor completely until the posterior part of the hard palate on both the sides so that we know the nasal mucosa is protected well.

I: Ok, when you raise the nasal floor completely, how do you it's completely raised?

P: [Picks up a skull and shows interviewer the parts of the skull associated with nasal mucosa] [this section is omitted as there is no context]

[Continued]

P: And another time I see trainees make a mistake is, when we are separating the nasal septum, most of them they lose orientation here [forgot where he said]. Because anteriorly we are at the septum, which is very easy to separate but as you go posteriorly we have the vomer bone here. I see the trainees tend to follow the septum and tend to ri de posteriorly superiorly, because the bony resistance is [word]. So, what happens is, this Osteotome, or this advancing Osteotome can go in there superiorly and can go into the cranium.

I: So, they are no realising it is going up.

P: No they are no realising it is going up. That is probably the main reason why the optic chiasm can get injured.

I: So how do you stop that?

P: I make it a point to tell them that when I am doing the dysjunction like this, I tell them that you look at the curvature of your Septal Osteotome and I ask them to feel the bone over here. And once they feel the bone I tell them to cut through the bone but not to go along the bone, so cut through the bone and follow the maxillary crest for a distance of around 3 centimetres. And of course, the assistant's finger is going to be there in the palate and the assistant will feel the advancing edge of the Osteotome and once he feels the Osteotome on his finger then we have separated it. That way I make sure they don't ride and go up. That's one thing you should do before downfracture. One is the dysjunction, using the finger and seeing that it is placed [word] Osteotome, and then cut through this and go across over here. Or sometimes I place an oscillating saw and weaken this. The third point I mentioned was that in one case, a couple of cases, I was not able to cut it so I did the Osteotomy after the osteotomy line was marked and then I cut this bone [REFERENCE] along with the maxilla and downfractured.

I: Yes, so I know the check for the dysjunction is to feel for the chisel to fully go through, but in cases it has not fully gone through, so only a part of the chisel has gone through.

P: Yes

I: Am I right in saying a check is to make sure the whole chisel has gone through?
P: Yes, to make sure the whole chisel has gone through, and then we achieve a complete separation from the maxillary tuberosity.

I: But you can't see it?

P: You can't see it, it's just feel. That is one of the reasons why the trainees feel it's very difficult when it comes to this area. They don't understand what I'm doing they are not very sure. And after about 5 or 6 cases where I ask them to do it themselves, I ask them to put a finger inside and then I do the Osteotomy, I mean I use the Rowe's Osteotome. And then I ask them when im…

I: Ahh, that's tough that! [Laughs]

P: [Laughs] I got to go very slowly [Laughs], I got to go very slowly to train then on how much to go and what level. That's probability the reason you know I said about the double tap. So that they don't hit it very hard, so they keep a check on how much force they are putting on the chisel with the mallet, so that they will rarely tear through and come out the side, compromising the blood supply to the maxilla.

I: Has anything ever gone wrong? If so, how do you react if something goes wrong?

P: Erm, yes, I should say fortunately I've never had to abandon a surgery. Because if the, your chisel goes through and cuts off the blood supply of the site, then I know the blood supply to the maxilla is compromised. And then you have only this posterior flap which is the gingival supply [inaudible]

I: Yes so you would have to cancel the surgery?

P: You just have to cancel it, there is no other option.

I: Have you even had a situation where it's gone wrong that was unexpected?

P: A bleed yes, during dysjunction.

I: For me being a non-surgeon, I would just panic, but how would and do y Osteotome you react when something unexpected happens?

P: You see, the thing is, with experience you understand that when there is a bleed in an inaccessible area, that you can't visualize, there's not much option really apart from applying pressure.

I: So as soon as it bleed,

P: It bleed

I: But what do you do, do you tell others first, you see I don't know the procedure, what do you do?

P: Basically, tell the anaesthetist there is a bleed and erm, they say 'hold on' and they try to get the blood pressure down and do whatever is possible from their side. And then immediately call the assistant for a [roller] gauze pack to place inside so they can apply pressure in that area. The third thing what I immediately would do is I will ask the assistant or the resident to give me a mouth prop, so that I insert the mouth prop between the mandibular teeth and the maxillary teeth so that I can exert pressure externally so the pressure is directly conveyed to the posterior maxilla.

I: Right so you place a gauze pack in the site and then use a mouth prop to leverage external pressure onto the site by using the mouth prop.

P: Yes so that the mouth prop in addition, erm, the pressure which is conveyed externally through the mouth prop also the maxilla [something]

I: That's good. Now if that happened to a trainee or a novice, how do you think they would, or how do they, react? There is a 'panic' feel, how do you also deal with that?

P: erm, the thing is when that happened I would just tell the trainees whenever it has happened not to panic, and I just explain, the thing is what they should understand is, there is not really a lot you can do at that particular point to control the bleeding, there is no point trying to run around for a [mosquito (scissors) or an artery forceps or using a diathermy, or electrocautery and try to buzz and burn up all the area without identifying what is the source of bleeding. Because you don't know whether it could be coming from the buccal side, from the soft tissue, it could be coming from the vessel, it could be coming from the pterygoid plexus of veins or it could simply be coming from the herniated buccal part of fat. Or most of the times it so happens that when you downfracture, the bleeding which is accumulated on the lateral side could be coming from the descending palatine vessel- which is more medially located. So there are multiple areas that it can bleed from, so I would rather advise them to put a pack and hold on for 5 minutes…

I: you pack, and by doing that you…

P: think it over a little bit, and then in the meantime if you are fairly young surgeon who is performing surgery for the first time I would say that, before you take up a case something for an osteotomy, and elective case, I would say that you at least consult with one other surgeon who in case, if you need, should be available to you. Ok. So in that case you call for somebody and additional help which you will require, somebody who has worked in the field, in this area, maybe an ENT surgeon or another OMFS surgeon senior, or a college OMFS surgeon, so that 2 of you take stock of the situation again, and then remove the mouth prop, remove the gauze pack again and check, 95% of the time you will find out the bleeding has come down considerably. And then you will put a retractor, and you will see where there is a blood vessel which is bleeding from the buccal side, they can use a bipolar, and then cauterize it, and then you get the maxilla down and you look inside and find out the descending palatine branches are bleeding, under which you can control it, it's fairly easy once the pooling of blood is lesser there and your vision improves, you can control the bleeder much easier. If you're still not sure and there is still some residual pooling of blood going on and you're not sure what is happening, then you get surgicel, back up surgicel, place again the gauze back on it, wait for ten minutes.

I: so you never think to yourself what might have gone wrong

P: it's very difficult…

I: You never consider what has gone wrong? So, if something happens do you first think of what might have been the cause of the mistake?

P: First thing would be to control the bleeding.

I: The first thing to come to your mind is that?

P: "Control that", yes, get it under control. And then start thinking about it. Whether my osteotomy cut was higher, 'did I do the dysjunction correctly', when I was retracting the soft tissue, did my channel retractor or [biterwarder] was retractor pushed in inadvertent and have I 'injured the buccal part of fat?'

I: Right, why do you go through the processes of what could have gone wrong, why does it help you?

P: That will start…because...as I am controlling the bleeder, I will start thinking of which other area I need to address.

I: Right, so you are going back?

P: Yeah, back to go forward. So, I will start looking at probably the first and the easiest cause and the most immediate and most common, so first I open that are and whether that is sorted out. Then I will go on the medial side and I will check for the descending palatine. That's easier and very accessible. I would control that, so that, you know, two of the major sources which could have been the reason, you tackle those. Then you go and check for the one that is absolutely not visible.

I: [laughs] wow that's a lot, don't make that error!

P: [laughs] so we go from one side to the other and try to look for all the sources, then finally you will be able to cure the problem, the one which it the most likely source. And then you know what should be done. You pack it and most of the time, I would say in 100% of the cases whenever I had bleed. It's only the pressure, finally, pressure and electro cauterize almost all times. If nothing of that kind happens, then I wouldn't go for any [external ligature] 3140. It is better with the pack and position, to shift the patient to a…in the hospital we have a cardiac OT where other angiograph facilities are available, so give them [SOMETHING] suction and angiography, find out the bleeder and perhaps the last solutions is to embolize the bleeder. If none of these things work, then I still wouldn't try to just simply jump into anything, erm, I'd go angiography, and then I embolize.

I: I imagine that can be a high stress situation.

P: Uh, huh.

I: With trainees are you able to explain to them your decisions, or are you focused a lot.

P: With the trainees, yes, it's quite stressful to do surgery. We have to train then and they are not aware, they don't know where to retract, of what I expect from them, they have no idea. And what do I want to see, they have no idea. So, it takes for them to assist me on two or three cases to get the idea of what I expect from them.

I: It takes them a while to hand you things before you ask for them.

P: Exactly, and engage them. [Phone call interrupt]. So, it is quite high stress when you are working with the trainees, who have come down for the first couple of Osteotomies, and in that case, virtual training if they have had it, they have an idea as to what is going on, what the prof is going to expect next. So, they are in a better position to [phone rings again] sorry.

I: No, its fine sure take it. [Break 2 minute]

I: We mentioned there is a lot of things going on, and a lot of people. When something does happen is it better to talk a lot about the problem, a maximal amount of talk to try and solve the problem at hand, or is it better to talk a little, so minimally, the basic information to try and solve the problem at hand?

P: It would be much, see what happens is, when we are working, because we are doing a large volume of cases, so it so happens that, erm, the Operation Theatre nurses, the guy who is assisting, they already have an idea about what I ask for next, so they come prepared. So not much of a talk really happens.

I: If anything happens they know as well?

P: They know what I ask for and they get it immediately. So not much talk happens. {Shared situational awareness] If I had to do a same thing in a university theatre apart from this hospital, it would involve a lot of talk.

I: is that because they might be less experienced?

P: less experienced yes. Because they don't have an idea of what I'm asking so I will tell them, this is exactly what I want. So, in those situations where I'm operating and in a theatre apart from the university, erm, setup, therein I need to go first, and then I need to see that the equipment is kept ready, then I imagine a little further and then I tell them ok keep a roller gauze pack, and get the surgicel available to me.

I: so you're not just thinking of what you're going to do, but also what they need to do, even before you have started? Because they don't know...

P: No, so I go and you know…

I: how many people do you usually have to think ahead for, so if you did need to think ahead for somebody and take charge, is it usually one or two people or quite a few more?

P: erm, its almost always, it's the assisting, the nurse, the OT nurse, she is the one that I would direct most of the talk to, so I tell her, give this to me, and then I tell the surgical assistant, the resident to actually concentrate erm, at the surgical site. And I ask them to do the suctioning, adjust the lighting, and then I want to look at this particular point. I just talk to them and tell them this is what I want to do, I ask for electrocautery they have no idea {inaudible]

I: so two people, two main people.

P: two people.

I: and if they have had less experience you have to keep speaking and for two other you think ahead for.

P: yes exactly we have to think ahead.

[Phone rings]

P: Sorry, sorry, sorry, sorry {break 2 minutes]

P: Sorry I missed you last question because of the phone call.

I: No it's ok, in fact you have been really helpful and there is more than enough here

P: Not an issue at all, you can talk tomorrow again I've got no surgeries tomorrow.

I: [debriefing information about what the participant's data will be used for]

P: [starts talking about the topic again] it's not necessary to clutter up the patient with too many things at the same time, the good thing is to have a calm mind and think about whether the bleeding is coming from the soft tissue, bone, which vessel and then identifying one by one, and then you go across check, you will find out the problems and you will be able to tackle it much better.

I: [talking about surgical experience and how interviewer observer errors in trainees and the need for training]

P: [general agreement on the interviewer's research] the thing is, when things go well on an osteotomy is much easier to learn [laughs], if everything goes smoothly you are not bothered, you know, you look at that as a glamourous surgery. But when things go wrong, that's the moment you have to be mentally and physically both ready. And co-ordinate both the faculties to see if you can sort the problem out at that particular point, without causing any harm to the patient.

I: right thank you!

P: pleasure, pleasure

I: lets go back and see what they are up to.

[End of interview]

My notes and assessment if they have the developed NTS of an expert:

Says the most critical point is the ptd and the tuberosities.

1. Some cases the bone of the ptd is fused which makes it harder to crack,
2. Also, the orientation of the grater palatine vessel/artery
3. Also, visuals, no one can see the ptd

Finger in hammerous region palpate.

Venus plexus is behind PMJ this is whan you should avoid hitting.

Greater palatine artery is what you need to avoid when downfracturing

Solutions:

1. Put the reciprocating saw in the junction and go through it
2. Use a smaller Chisel to be more precise and careful, only then after you have done some of the work can you use the normal bigger Chisel to finish the dysjunction or the reduction of the bone round the GPV.
3. He lets the novicres hit the hammer and he hold the Chisel, the nurse palpates for the Chisel. And after 5-6 csses of novice holding the Chisel he lets them feel ptd and mkes them tell his to stop when they think he has placed the Chisel in the PTJ.

There is also the ootion of the 3rd molar where you remove it to make easier access, I still am unsure what they mean. But the sme did say that with 3rd molar method, the Chisel can slip back and go behind and into the socket. A way to stop this is the shark fin osteoma, where by there is a small metal pertrusion that guide you. Not sure this is relevant. Sme said he tells the novices to use 2 taps, like a beat tap.tp….tap.tap….tap.tap..

SME holds the Chisel and positions the Chisel angle. he doesn't let them hold the Chisel as they cant see the junction so may not have it in right place, and even if he positioned it then after even 1 mallet hit .The older the patient the harder the ptd because of bone change.

This is because he knows novices will compensate for just using one tap, and over exert force hitting the hammer , but telling them to use two taps somehow distributes the force over those two taps.

When removing bone from around the palatine vessel/artery, he visually identifies the ptv and carefully wedges an elevator between the bone and the ptv, he then can remove/reduce the bone without damaging ptv.Nasal mucosa = Elevate to expose intra orbital nerve.Sme said that when error is made he firstly packs if bleeding, and then thinks back where the most likely mistake was, but he is constantly aware of the current situation, he then comes to a conclusion to where error may be and then that helps his for future choice

SME said he goes back to most likely cause of error by visually picturing what he has just done to mentally review the actions, starting with where he thinks to error most likely occured. this means that some of his working memory is used by this process but he said that while he is doing this he is still concentrating on the error control. (i think this may reduce his projective situational awareness, but at the same time, if he doesn't know where the error lies, then he cant project this information ahead to what will happen next, so he kind of has to look back to establish an educated guess of where the error lies, to then project and predict what will happen next and how to solve the issue.

Communication:

he said that when he teaches he is communicative to trainees even when decisions are quite tough, this shows he has the ability to separate his focus to differing objectives but still maintain a significantly better performance. the consequence of this was that is is quite tiring for him to have to use so much mental resource as to be aware of so many things for so long.

If he is in a room with many good staff there is minimal talking as they all understand each other and can predict each other's actions and thought processes. However, when SME was asked about hypothetically if his staff was quite new he was aware that this would mean he would have to tell them what to do next as he understands they would not be able to think ahead. He implicitly knows they do not have level 3 situational awareness, but he knows they have either level 1 or level 2. The consequence of that is that he must not only be thinking about the procedure but also preparing them to help him next by telling/directing them what they need to do next and be prepared for it (side note, my work can help them think ahead therefore alleviate the SME's stress as they would be able to think ahead and at least be able to have the next tools ready and understand the SMES cognitions to try to help them during error).

Sme thinks ahead not only for himself but also for 2 more people, the scrub nurse but also the trainee.

how to check the nasal floor is separated completely:

when he is elevating the nasal mucosa he used an Osteotome to push back the mucosa and he knows how far to go back as the Osteotome feels a slip on the end of the Osteotome when he reaches the back of the nose, this is because the bone drops down a little, so he felt less,a= like a give in pressure and sliding.

when novices separate the nasal septum the error novices make is that the novices start to hit the double/single (i forget) Osteotome with the mallet at the correct angle however the angle of the Osteotome moves upwards as the bone curvature makes this occur, the novice may not notice this angle change so that when they hit next the Osteotome goes further up and separates the bone too high, therefore they can even go through into the optic and do loads of damage. The sme tells the novice to Chisel a small chip into the nasal septum, this then makes the Chisel stick into the correct angle , then when the novice mallets the Chisel the Chisel has been set up to be able to go to the right angle as there is no other way it can really go.

Follow crest , assistant can feel the edge

first things he does if there is an error in the pterygomatic dysfunction:

tell the anaesthetict there is bleeding, this infers (as they both have experience and have worked together) that the anaestheseologist should now try to reduce the blood pressure to reduce the bleeding. she then asks For gauss strips from someone, to then start packing. this is more like a reaction than a startle effect, its like over time he has learned that this startle should be immediately replaces by this reaction. Sme holds on for time by packing and because of these two damage control actions (anaesthetist and packing) then he now can think about the best next step. This

is where he says novices differ, if youre a novice with an error he says you should call someone from another room in, as a back up plan. This is because it can go wrong fast so good to have an sme in the room also. But if you are an SME then you can continue with trying to stop the bleeding.

There is an importance for the visual, once bleeding starts, if there is too much bleeding and the visual decreases then is becomes vey difficult, this means you have a very short reaction window to pack and reduce blood pressure. If not then you have to again change plan to stop bleeding without knowing where it is coming from.

(sme actually opens to mandible with a mouth prop and packs the gauss strips into the ptj and then closes the patients jaw to apply their own pressure along with the patients jaw pressure.

there are many places bleeding could be from, says the SME, it may not be one near the ptj

is could be from buccal, or soft tissue or other vessles.

Sme says the downfracture is fine, as long as the checks are done beforehand.

He also sayd the ptd is easy to recheck as you put the Chisel back in the junction, this is easy to do as there should be already a crack/hole from the first attempt. Then put the Chisel in and mallet until assistance actually feels the full Chisel. Second attempt is easier than first attempt.

Pterygo-Maxillary Disjunction:

Situational Awareness:

- ‘feel the posterior dysjunction, that is the tuberosity cut at the Pterygo-Maxillary dysjunction, I feel is number 1, most other trainees find it most difficult to understand’
- ‘I: What do you find most difficult? P: [5 seconds pause] the dysjunction.’
- ‘And 2, the orientation of the Inter-Maxillary vessel is fairly closed’ (talking about why PTD Is critical)

- ‘you are going to operate in this particular area almost blind when you’re doing the dysjunction. Because you can’t visualise where exactly you’re doing the dysjunction most of it is tactile sensation because you want to keep one finger intra-orally in the hammerous region, you’re placing the Chisel externally and the third person is going to, you know use the Osteotome and, you know, hammer it. So, it’s quite a blind procedure.’
- ‘I can place the Chisel but I cannot see the end. I: So, no one can see it? P: No one can see it. It’s very difficult to see exactly where the Chisel is going, its mostly tactile sensation.’

Creativity:

- Generally, we have different sizes of osteotomies, and this is what I do, I’m just telling you, traditionally, in some of the cases where I am really worried I use an oscillating saw and I put it in the dysjunction area and I run it a couple of times. So, that the area becomes slightly weakened.
- In the junction, yes, just behind the tuberosity, where the tuberosity ends, I put oscillating saw here and then I just cut it so that it becomes slightly weakened. And second is, the other size Chisel is a bigger one, whereas there is a smaller size available in the Pterygoid Chisel, so I use a smaller one to work my way through in the dysjunction.
- P: And then the bigger one so that I know that the direction that I want the Chisel to go is going in the same direction.
- And I’ve also tried a few cases where the third molar tooth is removed and I can go through the socket.
- called a ‘shark fin’ Osteotome, it has got an additional projection, which, when you do the Le Fort 1 osteotomy cut you put the shark fin osteotomy line and then you give the angulation and tap it. So, that fin will kind of guide you.
- I think it is the oscillating saw and working with a smaller Osteotome and then increasing it a little bit, is the one which works the best.

Decision-Making:

- P: First thing would be to control the bleeding.

Question: Are the Pterygoid plate and the Maxilla separated?
Checks to answer question: Make sure the chisel has travelled through the dysjunction site.
How to do this: With a finger behind the junction, the chisel end should go through the bone and the finger should feel all the chisel end on the finger.
Method to accomplish this: Use a Mallet to hit a chisel through the junction until the chisel is felt on the finger
Decision point A: Has the Chisel end gone through the junction?
Decision Options: Fully, Partially, None.

Decision Issue Variables: Lack of Visual Feedback (I: Am I right in saying a check is to make sure the whole chisel has gone through? P: Yes, to make sure the whole chisel has gone through, and then we achieve a complete separation from the Maxillary tuberosity. I: But you can’t see it? P: You can’t see it, it’s just feel. That is one of the reasons why the trainees feel it’s very difficult when it comes to this area. They don’t understand what I’m doing they are not very sure.)

Decision Issue 1: If the chisel is not through enough, this can create later complication.
Decision Issue 2: If the Chisel is pushed in too far, after bony separation then this can cause sudden complications.

Checks to ensure decision issues 1 and 2 are prevented – Use the Mallet to hit the Chisel handle with an appropriate amount of force(s) so that the chisel will be controlled sufficiently to fully contact the finger, without going too far, or too short into the Pterygo-Maxillary junction.

Expert evidence to support this: 'Erm, in some cases its 3 or 4 taps, in some cases it might be 2 taps.', P: It's better to do a smaller tap, it's more controlled, and the tap is not a single one, it is two taps.

P: The ones in the tap, generally what happens is. I see the trainees when they put one single tap they tend to use a lot of force. You know, so it goes a little uncontrolled. So, when I tell them to do two taps so they tend to use the force in a more duratious manner. So, that probably is the reason why I tell them to do two taps.

I: Yes, I get you, to distribute the force

P: It distributes the force a little, not slowly but in a more definite manner. One single tap is generally more muscle and they don't tend to gauge how much force is needed. So, 2 taps probably, that's how I teach it.'

Figure 8D: Section of interview transcript of a non-SME interview

I: How many times did you shadow or assist before you performed your first Le Fort 1?

P: I don't exactly know how many I assisted, but roughly probably 25.

I: Let say less than 50?

P: Yes, less than 50.

I: Right ok, so did you feel prepared when you performed your first one?

P: Erm [pause 2 seconds] somewhat.

I: So, a little more and you would have felt prepared?

P: Yes, I was prepared but still once you get to do it on your once then it's a confidence thing.

I: Right. Now you're really confident and I imagine you train others?

P: Yes

I: Do you train them quite a lot?

P: Not so much once in a while I do.

I: Yes ok, so when you do a Le Fort 1 procedure, what do you find most difficult to perform, so what is the hardest part of a Le Fort 1, in general? [Pause 2 seconds]. It doesn't have to be a specific part, but just a part you feel is the most difficult to perform.

P: Erm, the posterior, when you to a pterygoid dysjunction. It's identifying the Greater Palatine. Sometimes trying to preserve the Greater Palatine becomes a big difficult part.

I: The Artery?

P: Yeah, the Greater Palatine Artery.

I: Right, so why is that difficult?

P: Sometimes, you may not be able to see anything and when you downfracture it might start bleeding. You try to preserve it any identify it, slightly difficult, not very difficult, but still difficult.

I: Out of all the Le Fort 1 that is the most difficult?

P: Yes, even you have to spend some time, there is a bone around it, you have to remove the bone, identify it and… see here [refers to a skull] the Greater Palatine Artery comes here, comes to there and supplies the palatal mucosa. So, when you downfracture that it'll be higher up here. So, when you downfracture it becomes difficult, other than that it is not a problem.

I: How do you prevent any errors? How do you stop errors from happening?

P: I take time that's what most important, I take time and do it cautiously.

I: But what do you do, what do you do to perform a dysjunction?

[Participant was describing a downfracture but the previous conversation was regarding a Pterygo-Maxillary dysjunction, for some reason they switched to describing this other task]

P: Dysjunction, just use the spreader [spreader can only be used for a downfracture, although creativity may accommodate a new use, the tool itself is not able to physically be used for this task]. I use the Smith's Spreader that's all.

I: This is on the downfracture?

P: Yes, downfracture, open it gently and identify the vessels, and just gradually mobilise it.

I: Is it quite easy to identify the vessels? So, what do you look for?

P: On cleft patients, it is difficult. Conventional osteotomies that are not cleft patients is easy, but in cleft patients it is difficult.

I: Ok. So, let say as a trainee, you are learning this, was it easy then to identify the Artery? Or did you perhaps struggle at first?

P: Post doc in most cases it is obvious, but 50/50. When I started doing it on my own then I got more confident.

I: How did you learn to identify the Artery?

P: You know the things with the anatomy?

I: [nods]

P: So, I know the anatomy, it's not just identifying it's trying to preserve that. Sometimes you have a thick bone around that, trying to do remove that small piece of bone, that's the time consuming and bit difficult part.

I: Yes. So, this is new to me, when you say preserve what do you mean by preserve the Artery?

P: The Artery is... once you downfracture, the jaw is opened up and the Artery is visible to you, sometimes there is a bone surrounding that, ok.

I: Yes.

P: Ok, the bone sometimes is very tight, so you have to remove surrounding that.

I: What's the bone called?

P: Erm, around the Greater Palatine, you have a bone [couldn't recall bone name].

I: Ok, so that takes some time? You cut and remove the bone?

P: Yes.

I: And what tools do you use to remove the bone?

P: Erm, sometimes I use small Osteotomes, or sort of, you know, sort of Artery [inaudible], you know.

I: Is this quite a common thing, to remove the bone around the Artery?

P: Not common all the time, sometimes we leave it like that, but sometimes it has to be removed.

I: When do you decide whether it should be removed or not?

P: Yeah so, the bone is too thick around the Artery and its impeding the impaction. During impaction, it becomes very important.

I: Right ok. So, do you find it quite easy to find the Pterygo-Maxillary junction with the chisel?

P: I don't do it all the time… just with the spreader I get dysjunction, or once in a while I use that.

I: So, you don't do the dysjunction, you go straight to the downfracture?

P: Yeah, downfracture, plus sometimes it happens sometimes it doesn't happen. You take this thing and you do it.

I: Right so you can do the downfracture without?

P: Yes, it'll happen.

I: does it split differently then?

P: no, no, no, not... See posterior cuts are proper, then erm, pterygoid, erm this this [referring to skull] automatically just, put a little, this thing it happens.

I: Right, so you tell it where to go?

P: The posterior cuts, see if it is here, you go all the way up do it here [referring to an extended back cut on the Maxilla so there is less bone before the end of the buttress], then it becomes easy and comes down. If you are giving it here, then you have to go and cut the osteotomy [referring to a 'normal' cut that stops at the third molar, therefore leaving more bone, requiring a Pterygo-Maxillary dysjunction]. The cut has to be all the way posterior.

I: Ahh, I see. So, that is better than…

P: No not all the time, case to case, yes sometimes you have to do it, sometimes you don't have to do it, if it is not moving then you have to go and do it [the dysjunction he is referring to].

I: Right ok, so there is difficulty as you can tear the Greater Palatine Artery, and you try to avoid that. [Back to the downfracture, as the topic strayed from primary task to secondary task difficulty]

P: Yeah, not very difficult but once in a while you have to be more careful.

I: What other things can you do to protect the Artery, do you anything else, maybe at different points?

P: No, the thing is when you downfracture, you just visualise the Artery. Don't do it forcefully- spread it all completely with the spreader.

I: Yeah

P: Gradually you see it, to prevent damage, see the Maxilla, if you… the Artery is tight, constant and too much, sometimes if you do it forcefully and fast it can tear the Artery.

I: So, you can see it stretch?

P: If you just gradually open up, see the Artery and everything, and your cuts are proper and everything, then gradually, complete downfracture -slowly, and then visualize the Greater Palatine Artery. Other than doing it a stretch. See if you directly do it. Sometimes it can sever and cut it.

I: So, if you were doing it slowly, would it stretch?

P: It will stretch to some extent, and it's not only that, you have to gradually see whether the cuts are perfect or not. The posterior aspect, not the anterior.

I: Yeah, not the Maxillary cut

P: The posterior aspect you can't see. So, once you downfracture only then you can see from the inside.

I: Right, what other check do you do on the downfracture? Others I have talked to state that for a successful downfracture you have to check certain things before, could you expand on this?

P: Yeah sure, for the downfracture you make sure the septum is separated, also make sure the cuts on the lateral osteotomy and the pyriform aperture, make sure the cuts are extended all the way. All those things. Then use the Osteotome to see if it goes smoothly and if there is any resistance, sound, all those things. And once you get the indication that the cuts are all clear, then you can do it.

I: Did you say sound then?

P: Yes

I: What sound does it make?

P: How can I describe it? The sound is erm, when you hit [pause 1 second] sort of, a hollow thing the sound gets, you know less, I don't know how to describe it.

I: Is this for the nasal aperture?

P: The lateral pyriform aperture. So, when you go a certain length, can you get that bottle? See notice the sound here (hits bottle and produces a hollow sound), how do you describe that? So, when you are cutting the bone this sound is {bit????}, the sound is not generated much.

I: Ok.

P: So, once you finish the cut it's a very different sound, when you are cutting (hits bottle to produce a lower bass tone), I don't know how to describe this. It's really difficult.

I: And is that the same for the dysjunction too, with the noise of the chisel?

P: Dysjunction I don't do, not that much.

I: Right yes, I remember you saying, but how do you make sure it is separated before you do the downfracture?

P: Using Osteotomes.

I: You just go back...

P: Do you mean the nasal mucosa?

I: Yes the nasal mucosa, because when you downfracture the nasal mucosa can still be attached.

P: We do it differently, the nasal mucosa has all these things that are separated differently.

I: How do you check there is no mucosa attached before you downfracture

P: It should be done before downfracture, it's obvious because once you are raised it is not in contact with the bone. There is no obstruction because it is separated from the bone.

I: So the Osteotome can go straight down.

P: Yes so you put the instrument here and check it out.

I: What happens if you do start to downfracture and it is attached?

P: The mucosa tears.

I: What happens then?

P: You try to suture it back, put some stitches.

I: Does it bleed a lot if it tears?

P: There is a lot of bleeding only if the Greater Palatine Artery and also when you do the pterygoid dysjunction. Posteriorly with the vessels.

I: So the tear is easy to manage?

P: Yes, yes.

I: And have you experienced when it does tear and there is a lot of bleeding?

P: No it does not happen there.

I: Talking about the bleeding then, have you had an experience where something has gone wrong and there has been bleeding?

P: No, as orthognathic not very exceptional, some minor things.

I: were you surprised those minor things happened, were you thinking how did it happen?

P: no, no, the thing… an extraordinary exception not happened it is always minor things, which is expected.

I: So you prepare for the minor things.

P: Yes once bleeding and once I had to cortorise the Greater Palatine, that's all, it is not difficult. And erm, not very difficult, what I mean is, you just go and buzz that vessel, so then the bleeding stops. Identify this, it's a little time when that bleeds, as a trainee I had problems, like when I had to close and fix the thing, bleeding started so we had to go back and open it and we were wondering why it happened, and it was because there was fluctuation in the blood pressure.

I: Right. For me, if that happened I would not know what to do, when it first happens does it startle you in any way?

P: No, not very big. The only thing was I was concerned that we have to open again.

I: Right, so you're already thinking ahead?

P: Yes, again we have to go back and do the operation and check it.

I: Did you say the Greater Palatine Artery was hit.

P: Yes there was some bleeding from there.

I: When things like that go wrong, how do you tell the others? How would you communicate things like this?

P: Discuss with the anaesthetist just to go back and check again so we have to bring him in and ask him to reduce the pressure so that we can go and pick up the bleeder, put in a pack, or do something. [Words] because there is a bleeding and if necessary identify the donor of the transfusion.

I: If something does go wrong, is it better to talk a lot about the situation to try to solve the problem or is it better to try to keep the talk to a minimal amount in order to solve the problem?

P: Lots, when there is a problem you try to get to the problem, discuss with them, like if bleeding is there and you are not discussing it with the anaesthetist then you are doing a mistake. You try to discuss the reduction of the blood pressure or see the [word], and always then see if too much of bleeding, whether we need to do a blood transfusion or not.

I: And you decide that yourself or do you talk that out?

P: I decide best on the situation, most of the times it will not be required, but when there is a problem, yes you have to communicate with the people, you have to communicate.

I: Do you find that your team is supportive, for example if you need a tool they will have it ready.

P: It's based on the situation, but most of them yes, they know.

I: Based on the situation, what situations?

P: In a situation such as a panic situation, they tend to do mistakes. But otherwise most of the time it's normal.

I: Does that effect you if there is a panic situation?

P: To some extent, you have to do it fast…

I: For time pressure?

P: Yes time pressure, because of the bleeding we have to try to control the bleeding first.

I: Ah right, how do you manage? I was talking to someone before and they said they 'stepped back', but another person said they take charge and get right into it, how do you react and respond?

P: I put a pressure pack, wait for some time, ask the anaesthesiologist to control the blood pressure a little bit so there is a favourable BP so that we can go back and look into that.

I: Right so you buy more time to stop the blood.

P: Yes controlling blood pressure that is the best thing you can do for a bleed.

I: Right ok thank you for sharing that it is really useful. Is there anything else that you have found difficult to teach, not perhaps performing but to teach trainees?

P: [pausing to think] Sometimes giving cuts with the saw – for a trainee it will be difficult. So using the micro motor is easier, it is a slow process but it is safe. So the cuts with the saw, a trainee can go wrong because of the speed, and they might go in the wrong direction if you don't take control of them. That's the only thing I think that we should be careful when giving them to a trainee. Use a slow bur with a slow rpm. They can go slowly, if the speed is fast, suddenly they go posteriorly [something].

I: So you use the bur saw. Do you draw any marks on?

P: Yes, yes that's the best thing what you can ask the trainees to do, draw and gradually put some marks, and then join them.

I: Can you tell when they do go wrong…

P: We don't give them complete control, we only give it to ones who become a junior staff.

I: Ok, and have you ever see them before they go wrong and stop them just before they have gone wrong?

P: Most of my cases I do are clefts, so not really.

I: Right. Now, these questions are going to get a little stranger, but they have got a point for me within them. Ok. So when something goes wrong, what are the positives or negatives of thinking about what might have happened to make something go wrong. What are the reasons perhaps of trying to think back to what might have gone wrong?

P: It's so you can prevent the complication next time, so you can go back and look at that and think where did I go wrong, so you can think about it next time.

I: Is it good in the moment of the error?

P: No, no, no, no. See, the first thing you should do is what is important, which is the bleeding, that is the first. When you borrow some time then you can think where you've gone wrong, right. In the meantime. So then you don't think about that in the present condition in dicey condition.

I: Thank you it was kind of a trick question.

P: Yes.

I: Like you said you have to think later on what happened, but that shows expertise to think ahead, regardless of what happened. Have you ever heard of have had training in non- technical skills before?

P: No.

I: How important do you think they are? Thinks like teamwork, decision making, situational awareness, communication, and leadership.

P: Communication is very important, there might be seniors, juniors supporting me, it makes a lot of, see, and it gives a lot of confidence when you are doing surgery. You are going with the team. And you are clear of your concepts and you are taking everybody with you. Even if you are going wrong there is somebody to help you at that time when you're doing wrong, to support you at that time. So if you do that there are a lot of advantage, rather than standing in, trying to do that on your own.

I: What would happen if your team changes and it's not your usual team, what would happen? For example, imagine all your team is new.

P: I don't think it would make a significant difference. It's the experience that counts. If experience is there, I don't think it should make a significant difference.

I: Ok, and if they are newer, and have not had much experience?

P: It might take them time, the time taken may take them slightly more, might lose some patients [laughs] (he meant frequency not deaths)

I: And would you feel that you would have to do more?

P: More effort sure, more effort.

I: Does that affect your focus?

P: They know they might do mistakes and I would spend more time telling them. That's a bit erm, what do you call it, erm, unpleasant thing, to do it, at time when you are doing surgery you would rather concentrate on the work you have to do. You job is not to [word].

I: Right, ok, one more question? I don't want to take more time.

P: Ok

I: Ok, what are the other checks before the downfracture, do you do anymore?

P: The pyriform aperture area, the Osteotome put it in the centre and posterior.

[Can't tell what saying]

P: And sometimes Septum. Hopefully it was some use [we now talk about other stuff, I debrief him and tell him about how his data will be written up].

Figure 9: Debriefing Form for CAPs

Debriefing form

This study was concerned with your cognitions/reflections/experiences of a Le Fort 1 Osteotomy. Your participation helped to uncover the cognitive differences between experts (your role) and trainees, in terms of their technical and non-technical skills. This will help in creating a cognition-based training framework that will be implemented into an existing high-quality virtual surgical simulator created by PhD student Yeshwanth Pulijala at University of Huddersfield. This will ultimately create a tool that is equal to or better than current virtual and real-life training simulations.

<u>How was this tested?</u>

In this study, you were interview into your knowledge and experiences of a Le Fort 1 Osteotomy. This interview used a technique called 'Cognitive Task Analysis' to investigate your cognitions when you perform the procedure. This was not an evaluation or assessment of your abilities, and will not be passed on to external bodies or third parties. The exploratory interview simply investigated your thoughts/reflections/experiences of the Le Fort 1 Osteotomy.

<u>What data were collected?</u>

Your age, sex, and information on your surgical knowledge/abilities were recorded via Dictaphone and questionnaires. However, if there are any identifying (names, places) or signifying (history, hobbies) information, these will be extracted from the data to provide anonymity. All data is anonymous and will be destroyed 5 years after the end of the researcher's MSc course.

<u>What if you want to know more?</u>

If you are interested about learning more about this subject area, you can contact the following individuals:

Matthew Pears- Email Address: matthew.pears@hud.ac.uk

Dr David Peebles - Email Address: D.Peebles@hud.ac.uk

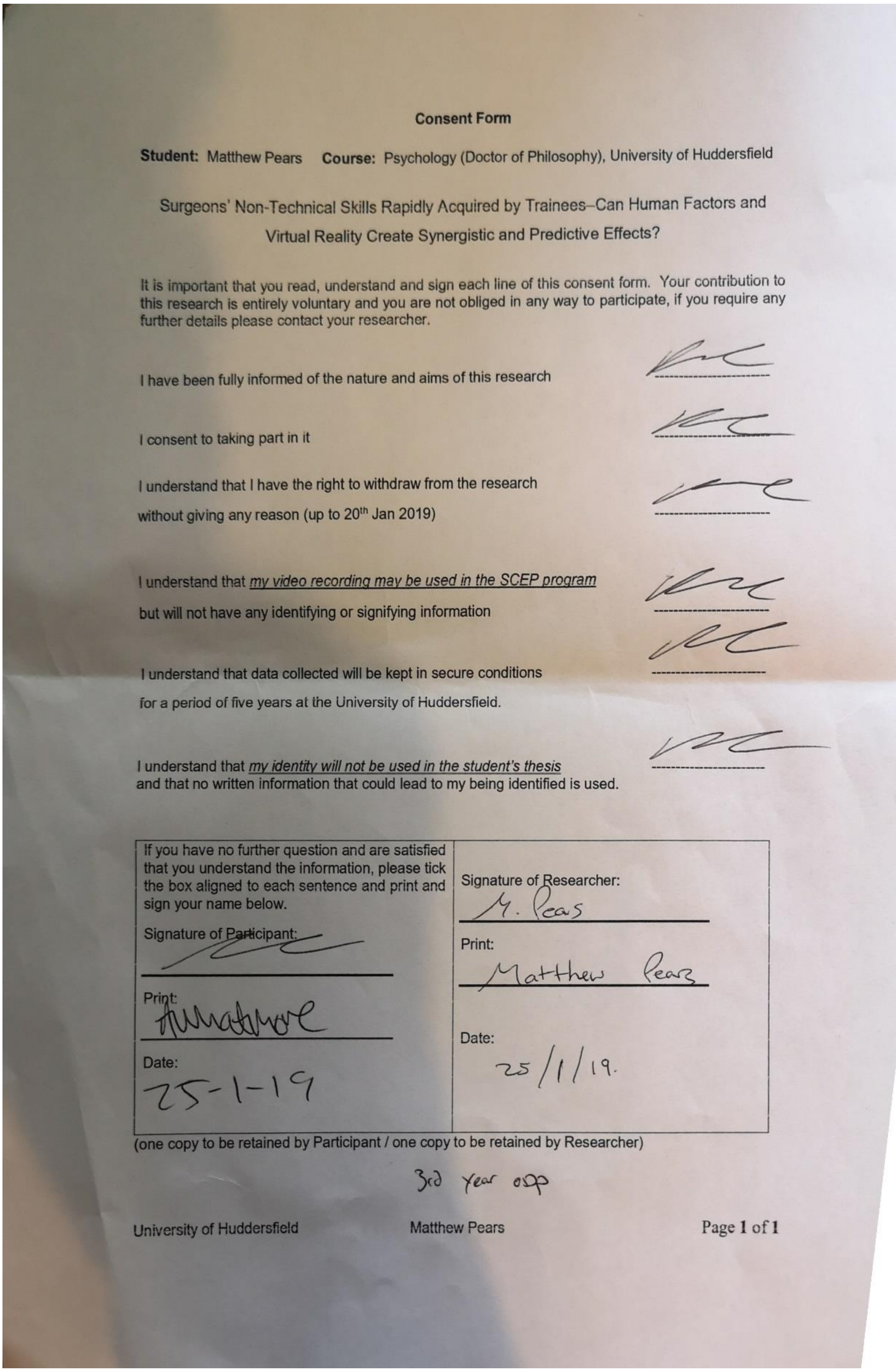
<u>What if there are any issues?</u>

If any issues arise, no matter how small, please contact the following individuals:

Matthew Pears - Tel: 07575151512. Email Address: matthew.pears@hud.ac.uk

Dr David Peebles - Tel: 01484 473173. Email Address: D.Peebles @hud.ac.uk

Figure 10: Completed Student Participant Consent Forms



Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research

I consent to taking part in it

I understand that I have the right to withdraw from the research without giving any reason (up to 20th Jan 2019)

I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information

I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield.

I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.	Signature of Researcher: M. Pears
Signature of Participant:	Print: Matthew Pears
Print: [illegible]	
Date: 25-1-19	Date: 25/1/19.

(one copy to be retained by Participant / one copy to be retained by Researcher)

3rd year osp

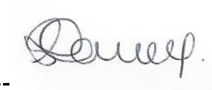
University of Huddersfield Matthew Pears Page 1 of 1

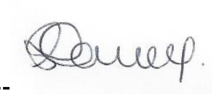
Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

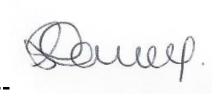
Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

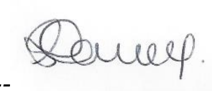
I have been fully informed of the nature and aims of this research -- 

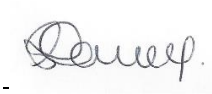
I consent to taking part in it -- 

I understand that I have the right to withdraw from the research

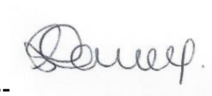
without giving any reason (up to December 4th, 2018) -- 

I understand that *<u>my video recording may be used in the SCEP program</u>*

but will not have any identifying or signifying information -- 

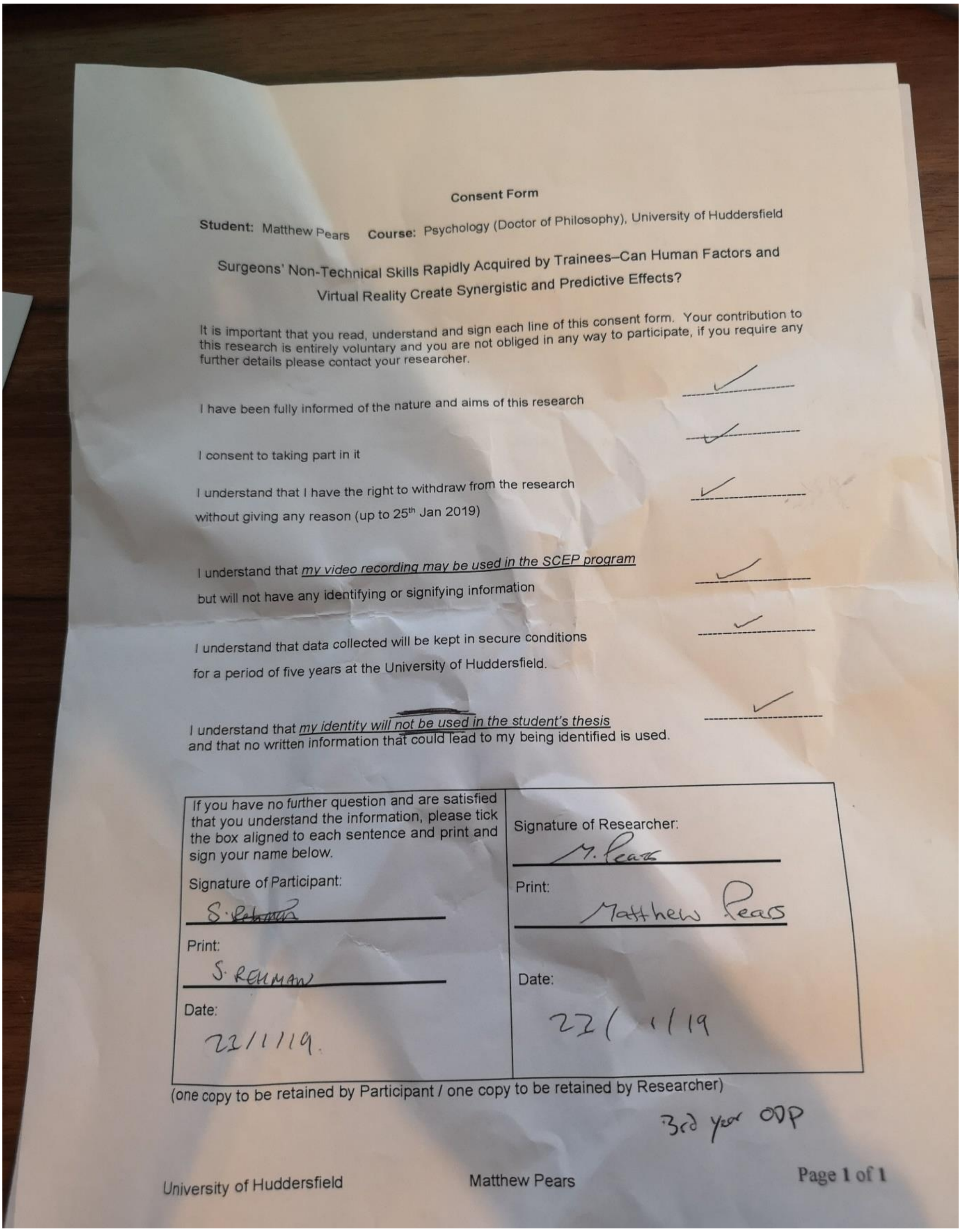
I understand that data collected will be kept in secure conditions -- 

for a period of five years at the University of Huddersfield.

I understand that *<u>my identity will not be used in the student's thesis</u>* --  and that no written information that could lead to my being identified is used.

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below. Signature of Participant:  Print: MISS C.DOWNEY Date: 09/01/2019	Signature of Researcher:  Print: Matthew pears Date: 09/01/2019

(one copy to be retained by Participant / one copy to be retained by Researcher)



Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	✓
I consent to taking part in it	✓
I understand that I have the right to withdraw from the research without giving any reason (up to 25th Jan 2019)	✓
I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information	✓
I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield.	✓
I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used.	✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.	Signature of Researcher: M. Pears
Signature of Participant: S. Rehman	Print: Matthew Pears
Print: S. REHMAN	Date: 22/1/19
Date: 22/1/19.	

(one copy to be retained by Participant / one copy to be retained by Researcher)

3rd year ODP

University of Huddersfield Matthew Pears Page 1 of 1

Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	✓
I consent to taking part in it	✓
I understand that I have the right to withdraw from the research without giving any reason (up to 20th Jan 2019)	✓
I understand that *<u>my video recording may be used in the SCEP program</u>* but will not have any identifying or signifying information	✓
I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield.	✓
I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used.	✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below. Signature of Participant: T.Taylor Print: Tanicha Taylor Date: 25/1/19.	Signature of Researcher: M.Pears Print: Matthew Pears Date: 25/1/19

(one copy to be retained by Participant / one copy to be retained by Researcher)

3rd Year ODP

University of Huddersfield Matthew Pears Page **1** of **1**

Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ______

I consent to taking part in it ______

I understand that I have the right to withdraw from the research without giving any reason (up to December 4th, 2018) ______

I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information ______

I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield. ______

I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used. ______

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.	Signature of Researcher: M. Pears
Signature of Participant: [illegible]	Print: Matthew Pears.
Print: SAMIYA HUSSAIN	Date: 3/12/18
Date: 3/12/18	

(one copy to be retained by Participant / one copy to be retained by Researcher)

Consent Form 10

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ✓

I consent to taking part in it ✓

I understand that I have the right to withdraw from the research without giving any reason (up to December 4th, 2018) ✓

I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information ✓

I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield. ✓

I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used. ✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.	Signature of Researcher: M. Pears
Signature of Participant: [illegible]	Print: Matthew Pears
Print: Zenib Hussain	
Date: 4/12/2018	Date: 4/12/2018

(one copy to be retained by Participant / one copy to be retained by Researcher)

University of Huddersfield Matthew Pears Page 1 of 1

Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	✓
I consent to taking part in it	✓
I understand that I have the right to withdraw from the research without giving any reason (up to December 4th, 2018)	✓
I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information	✓
I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield.	✓
I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used.	✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below. Signature of Participant: Print: Haris Mirza Date: 10/12/18	Signature of Researcher: M. Pears Print: matthew Pears Date: 10/12/18

(one copy to be retained by Participant / one copy to be retained by Researcher)

University of Huddersfield Matthew Pears Page 1 of 1

12

Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research	✓
I consent to taking part in it	✓
I understand that I have the right to withdraw from the research without giving any reason (~~up to December 4th, 2018~~)	✓
I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information	✓
I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield.	✓
I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used.	✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below. Signature of Participant: Print: H [illegible] Date: 10/12/18	Signature of Researcher: M Pears Print: Matthew Pears Date: 10/12/18

(one copy to be retained by Participant / one copy to be retained by Researcher)

University of Huddersfield Matthew Pears Page 1 of 1

13

Consent Form

Student: Matthew Pears **Course:** Psychology (Doctor of Philosophy), University of Huddersfield

Surgeons' Non-Technical Skills Rapidly Acquired by Trainees–Can Human Factors and Virtual Reality Create Synergistic and Predictive Effects?

It is important that you read, understand and sign each line of this consent form. Your contribution to this research is entirely voluntary and you are not obliged in any way to participate, if you require any further details please contact your researcher.

I have been fully informed of the nature and aims of this research ✓

I consent to taking part in it ✓

I understand that I have the right to withdraw from the research without giving any reason (up to December 4th, 2018) ✓

I understand that *my video recording may be used in the SCEP program* but will not have any identifying or signifying information ✓

I understand that data collected will be kept in secure conditions for a period of five years at the University of Huddersfield. ✓

I understand that *my identity will not be used in the student's thesis* and that no written information that could lead to my being identified is used. ✓

If you have no further question and are satisfied that you understand the information, please tick the box aligned to each sentence and print and sign your name below.	Signature of Researcher: M. Pears
Signature of Participant: [signature]	Print: Matthew Pears
Print: Mariam Tanvir	Date: 10 / 12 / 18
Date: 10/12/18	

(one copy to be retained by Participant / one copy to be retained by Researcher)

University of Huddersfield Matthew Pears Page 1 of 1

Figure 11: Patient Consent for Surgery Recording (NHS Clinical Media Glasgow has the single completed hard copy, available upon request)

The Royal Live pool and
Broadgreen University Hospitals
NHS Trust

Consent to Clinical
Photography Forms

Please ensure you have read and are familiar with the Trust's Clinical Photography Policy prior to performing Clinical Photography or consenting patients for same.

Page 1 Patient Information Leaflet

Page 2 Medical Illustration Services request (if required)

Page 3 Consent Form - copy for case notes

Ensure implications of CL4 are fully explained to your patient

Page 4 Consent Form Yellow Copy - to be given to patient

PATIENT INFORMATION

Consenting to Clinical Photography or Video recording

The Royal Liverpool and Broadgreen University Hospitals NHS Trust has a policy to give you the right to control the use of photographs or video recordings, which may be taken during the course of your treatment.

You can refuse to have photographs or videos taken for any reason other than for your health records. This will not affect your treatment in any way.

You have been asked to have medical photographs or video recordings taken. These will be for:

1. Your health record - you may not be asked for your written consent for this.
2. The teaching of health professionals and students studying healthcare here and in other hospitals/colleges/universities.
3. The education of patients with conditions similar to your own.
4. For publication in Medical and Scientific Journals or Textbooks either now or at any time in the future or for some other specific use that will be explained on the consent form.

You will be given information about what the recordings will be used for in numbers two, three and four above, and will be asked to sign a consent form.

You can say yes to as many or as few of the above as you wish. Please be aware that once photographs have been published, you cannot withdraw your consent.

Further Information

If you have any further questions please speak to your doctor or nurse.

Author: Clinical Photography Sub-Group
Date: July 2006
Review Date: August 2008

This leaflet is available in large print, computer disc, Braille, audiocassette and other languages on request.

Communications and Marketing system job number

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Complete only if requesting Clinical photography team to perform Clinical Photography

Please PRINT using **BLACK** ink. Give / Send to Clinical photography, Ext 2880 / 2886 / 2890

Requested by

Name ....

Department ....

Directorate ....

Contact No .... Date ..../......./......

Signature ....

Patient Details / ID Sticker here

Surname/family name...

Forenames ....

Date of Birth ....

NHS number ....

Unit number ....

Male ☐ Female ☐

Media format

Colour Print ☐ No. of copies ....

Digital Image ☐ No. of copies ....

Video Recording ☐ contact Medical Illustration to discuss

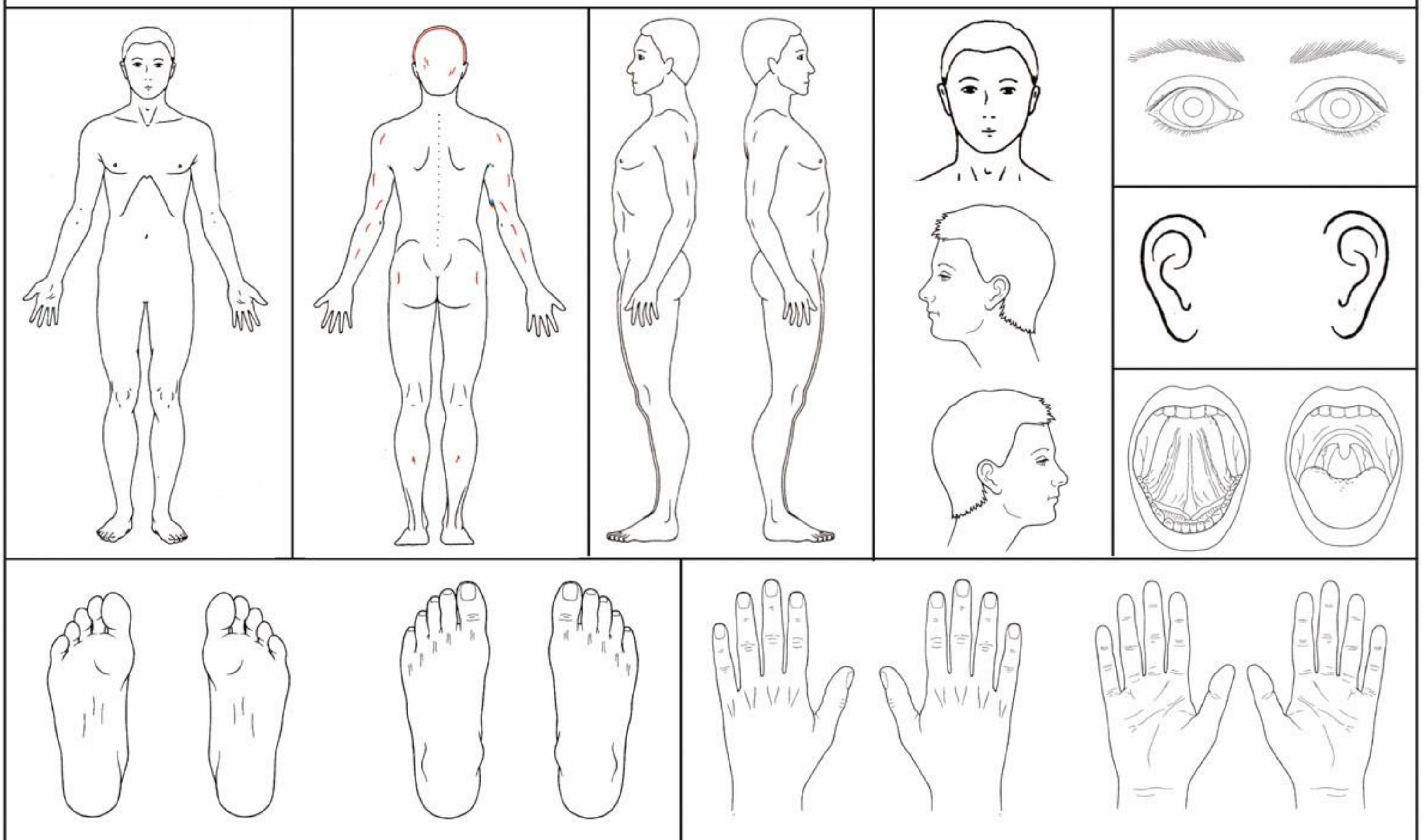
Consent Level

Pleast tick each level agreed with the patient as stated on Clinical Photography Consent form

CL1 ☐ **CL2** ☐

CL3 ☐ **CL4** ☐

Please outline views / areas required



Is a scale / measure required on close ups Yes ☐ No ☐

For Clinical photography use only

Location of Patient photographed:

RLUH	☐	**BGH**	☐
Studio	☐	**Ward**	☐
Clinic	☐	**Theatre**	☐

Date requested ...../...../...... Date photographed ..../...../......

Date completed ...../...../...... Ward / Clinic ....

Clinical Photographer ....

Signature ....

Consent to Clinical Photography/Video and Transmission form Please **print** using **Black** ink

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Clinical photography/video/transmissions to be undertake by:

Medical Illustration Services □□St. Paul's Imaging Unit
□

Trust registered health professional □□

Mersey School of Endoscopy

Anatomical area...

Patient Details / ID Sticker here

Surname/family name...................................... Fo
....

Date of Birth ....

NHS / Unit No..

I have explained the purpose of clinical photography/recordings to the patient and how the images will be used. □□□□

Patient information leaflet has been given.

I am a health professional requesting clinical photography/recording.

I am a Trust registered health professional performing clinical photography/video recordings/ or transmissions □

I will ensure that the appropriate images are taken in a dignified manner using approved equipment in accordan□□□ with Trust policy.

I will ensure all images used for the purpose of CL4 will **not** identify the patient.

Signature of health professional.. **Print Name** ....

Job Title .... **Contact details**.............................. **Date**.......... / .... / ....

Patient statement (please circle your answer)

I agree to have clinical photographs/video recordings/transmissions done. The request for the same has been explained to me and I fully understand what it entails. **Yes No**

CL1. I consent to clinical photographs/recordings being taken for my personal health record only. **Yes No**

CL2. I consent to clinical photographs/recordings being available for teaching in the health care context. **Yes No**

CL3. I consent to my clinical photographs/recordings being used to educate patients undergoing similar treatment w the Royal Liverpool and Broadgreen University Hospitals NHS Trust **Yes No**

CL4. I consent to my clinical photographs being published for the specific purpose of .... and/or* publication in Medical or Scientific Journal or Textbook at any time in the future. (* please delete as appropriate) **Yes No**

Signature of patient / parent / guardian* .... **Date** ..../......../........

*** Must have parental responsibility for the child**

Relationship to child ....

A witness should sign below if the patient is unable to sign but has indicated his or her consent

Signature .... **Name (print)**...................................... **Date** ..../......../........

Statement of Interpreter **Yes** ☐ **No** ☐ **Not applicable** ☐

N.B if telephone interpretation used, health professional to enter details below

I have interpreted the above information to the patient to the best of my ability and in a way which I believe he or she can understand.

Interpreter's signature ....**Name (print)**......................................**Date** ..../......../........

Yellow copy accepted by patient (please circle answer) **Yes** **No**

Consent to Clinical Photography/Video and Transmission form Please **print** using **Black** ink

The Royal Liverpool and Broadgreen University Hospitals NHS Trust

Clinical photography/video/transmissions to be undertake by:	**Patient Details / ID Sticker here**
Medical Illustration Services St. Paul's Imaging Unit Trust registered health professional Mersey School of Endoscopy **Anatomical area**...	Surname/family name...................................... Fo .... Date of Birth .... NHS / Unit No..

I have explained the purpose of clinical photography/recordings to the patient and how the images will be used.

Patient information leaflet has been given.

I am a health professional requesting clinical photography/recording.

I am a Trust registered health professional performing clinical photography/video recordings/ or transmissions

I will ensure that the appropriate images are taken in a dignified manner using approved equipment in accordan with Trust policy.

I will ensure all images used for the purpose of CL4 will **not** identify the patient.

Signature of health professional.. **Print Name** ....

Job Title .... **Contact details**............................... **Date**.......... / .... / ....

Patient statement (please circle your answer)

I agree to have clinical photographs/video recordings/transmissions done. The request for the same has been explained to me and I fully understand what it entails. **Yes No**

CL1. I consent to clinical photographs/recordings being taken for my personal health record only. **Yes No**

CL2. I consent to clinical photographs/recordings being available for teaching in the health care context. **Yes No**

CL3. I consent to my clinical photographs/recordings being used to educate patients undergoing similar treatment w the Royal Liverpool and Broadgreen University Hospitals NHS Trust **Yes No**

CL4. I consent to my clinical photographs being published for the specific purpose of .... and/or* publication in Medical or Scientific Journal or Textbook at any time in the future. (* please delete as appropriate) **Yes No**

Signature of patient / parent / guardian* .... **Date** ..../......../........

*** Must have parental responsibility for the child**

Relationship to child ....

A witness should sign below if the patient is unable to sign but has indicated his or her consent

Signature .... **Name (print)**....................................... **Date** ..../......../........

Statement of Interpreter **Yes** ☐ **No** ☐ **Not applicable** ☐

N.B if telephone interpretation used, health professional to enter details below

I have interpreted the above information to the patient to the best of my ability and in a way which I believe he or she can understand.

Interpreter's signature ....**Name (print)**.......................................**Date** ..../......../........

Yellow copy accepted by patient (please circle answer) **Yes** **No**

Figure 12A: Participants Coded Verbal Protocol Report Transcripts Before Intervention

Participant 1 VPR

The anesthetists don't look like they are paying attention to what's going on, and the others are sitting around talking. At least they are not on their phones, I guess. She's standing really close, that makes you really nervous. The music seems really loud that they have got on. Ah, you shouldn't pass…you shouldn't have a knife like that, you shouldn't wipe a knife, it should be in a dish. It doesn't look like they are ventilating properly. So I'm not really saying much as there is a lot going on its too fast. The anesthetist is glancing around so he is looking at the screen. She hopped over the tubes there she could have pulled them out. Is someone behind me, no, it's noisy I can't tell the direction of talking.

P2 Pre-Video Assessment Transcript

Participant:

I'm going to deliberately look around. There's something going on in the background with the runners. Pausing to teach. Continuing the operation. Requesting information. More conversations over there. What was happening to the left? [Laughs]. So I didn't really speak during the video but just pausing at the end yeah I should have paused it I didn't think to pause. The surgeon was good to be clear to the others what he was doing. I know there were several things going on around me but I didn't really look as I was anticipating something surrounding the surgeons. And the runners now talking in the last bit of the video, and the way they are interacting is making them to drop a few things they are on autopilot really.

Participant 4 Pre video assessment

Right so, two anesthetists. Quite a few observers. I mean I know people are talking and its busy, I'm feeling aware of the noise I don't know if its on all the time or not but with the talking then everything does add up so I would be conscious of that if I was really there, or at least that I wasn't part of making the noise. So the girl is close to the patient, nothing happening behind me I just assumed there wasn't anything as I'm usually just looking at the scrub practitioners area. So the surgeon is asking the runner and he has to repeat himself, I've kind of pieced together the noise is making the communication harder. I still think the surgeon is trying to get the runner involved to stop the talking but in a nice way. The scrub nurse doesn't talk to the runners at all, so I don't think they know where they are at in the procedure, one lady is looking over though. The scrub practitioner is really trying to keep up she's good at multitasking.

Participant 7e Pre video assessment

Woah this is different, it looks a bit flat but I can see everything. The anesthetists there talking and the monitor, the surgery is a head surgery but I cant see what it is. So there are quite a few people in here im guessing it's a complex surgery cos there are 2 anesthetists and 3 runners helping unless its swapping staff or something. The scrub nurse seems to be busy and is getting things and trying to keep up with the surgeon. The girl is really close to the patient she's not in green she should step back a bit because she might make it not sterile. The surgeons asked someone over, he's making her look at something maybe just so they aren't talking anymore. Its quite calm and I feel I'm just stood here; it feels weird as I would be moving. She's stepped over the wires there. The surgeon is good at telling people and showing the trainees.

8c Pre-Video assessment

So one thing I've kind of immediately noticed is that there's quite a lot of people in that theatre and normally they try keep it as like they try and keep it as low as possible I mean these might be medical students who are watching the procedure that we talked about, there's another person here and then there's three people sat there, they might be the circulators, but again they're not really fully paying attention to the procedure I mean they could be listening at the back kind of. But you don't know and again there's two guys over here, could be anaesthetists, could not be I mean one thing is do they need to be there? If it can occur in the anaesthetic room or it can be done later or it just like over lunch, then that would be more beneficial to the patient because obviously its less risk of bringing stuff in and stuff like that so that's the first bi . OK so I get that they've got to learn about the procedure but leaning over the drapes like that they're sterile drapes green yeah they're mean to stay sterile so then kind of leaning over and going kind of getting into the surgeon's kind of space there's a risk there is a risk of desterilisation there. It's happened to me before where I've been scrubbed and someone's stood directly behind me, I've not noticed I've stepped back and immediately walked into them and I had to change my entire gown and entire gloves. So stuff like that I mean I've done it where they've been more at the head end so there wasn't that risk but because they're just leaning over there is that risk of desterilisation and so again another risk of infection. Not really paying attention to the procedure yeah like I said they could be listening outside the back and it can drag a bit when you're sat not doing anything but you know they literally look like they're not paying any attention so when he spoke to the lady in the middle for the first time she had to ask him to repeat it. I mean it's happened to me before where I've not heard what they've said but you know and I was talking to someone about something and that was my fault and I admitted that and erm I've learnt from that and again and that is something that I think is being displayed here, they're just kind of chatting and may not be fully listening to what's going on and stuff like that.

Participant 10 pre-condition measure

So, the scrub nurse leans onto the patient, that's not good. ~~So, the surgeon calls in the scrub nurse to have a look~~ and she leans over the patient using her elbow which could obviously affect pressure areas after. ~~The surgeon calls over the lady on the right to sort of get~~ involved but she's sat down so she looks over at him and she comes over when he calls her. She's only come over because the surgeon's called her, otherwise she wasn't really bothered. She walks into the sterile field trying to avoid anything else that's sterile but it's obviously a sterile field for infection control, she shouldn't be in there. The gentleman on the left, they are talking away, guessing that's the anaesthetist and the ODP they're just talking away, they're not really focusing on what's going on, the anaesthetist just kind of feels that his job is to just look at the monitor, nothing else. So, the surgeon is trying to get that lady involved a little bit more, she comes over, has a look and goes back and sits back down and doesn't really bother. The scrub nurse hands over the blade to the surgeon using her hand, ideally, she should be handing it over in a kidney dish. And then she puts it back on the table without a kidney dish. Again, that's dangerous, anyone could get cut with that, especially the scrub nurse could get cut with that. She uses a dirty swab to clean the knife. The lady on the left is kind of leaning over to have a look inside the patient to see what the surgeons are doing. She could fall over; she could possibly fall onto the patient. The way that she is stood, her posture isn't very good, the way she is stood is not very stable. There's the suction pipe behind the lady and she attempts to hop over it. Again, that could be a trip hazard for anyone,.

Participant 14 pre-condition measure:

So I'm at the end of the patient, this is a different place than I would usually stand but I'm not doing anything so I can see everyone. Two anaesthetists. This is a normal room size it feels quite small I don't know if it's the headset or the people in the room, but it feels full. There are 3 runners sat down so there is maybe one more than normal, so I don't know, in fact the anaestnetisis are 2, usually only one so maybe they are having a swap over. I'm not sure I can't quite hear what they are saying. The scrub nurse is pretty busy. Oh, what's that the surgeons making the runne have a look I don't know why she's not a surgeon, but I suppose he's letting her learn about the procedure, she wasn't really interested but its good to share. So, the scrub nurse passed the knife by hand. The observer just hopped over the suction there I mean that's fine but its that what if moment if she catches it and the tube is pulled out. No-ones noticed or said anything, I think everyone are quite, kept to themselves, like just doing their own thing. So I cant see any major problem, just little bits and bobs of things.

Figure 12B: During Intervention

Participant 1 Experimental Condition Completed

Oh, it feels like I'm going to fall over. Oh, what's that? You shouldn't have your bag in theatre. Shouldn't have a knife. What is she doing? Oh no, she's putting all the swabs in a bowl, you should count them. There's no… anaesthetist, someone should be watching them. She should be watching her count the swabs out. You shouldn't just open a tray.

Part 2:

Option 1: That he has noticed that they are tired, I guess. There is nothing you can do- there is nothing he can do, there is no point in rushing cos it's a patient.

Option 2: Well, everyone is an adult, they can look after themselves really There is nothing he can do. You can't stop half way through an operation. If you're at work, you're at work. But that's a risk as you're trusting others to look after themselves, I suppose if you know them, then you over time trust their judgment, so only if I know them I would do this but it's only if the feel at the time is right

Option 3: I don't really think there is any benefit of pausing, cos then you're just taking longer than you would have been doing. The other staff can give each other a little motivational speech, I don't know. But you just need to take a bit more care if you're tired.

Participant 2 Experimental Condition

I see the surgeon is communicating with the scrub nurse. And the observer is causing distractions by asking questions. the scrub nurse is anticipating the next need of the surgeon. Ooh hello they er the runner has just put something in the scrub nurse's tray without her knowing. And is er writing it on the whiteboard without acknowledging anyone else there's room for error. See the nurse she's counting she's probably not aware of the other [incoherent] on her tray. The surgeon's waiting for equipment which means the scrub nurse is having to be rushed. She's leaving equipment on the drapes which could then get lost. She's also moving away to assist the anaesthetist which is a weird thing to do. There's tools here hidden underneath the swabs. Its quite clear communication from the surgeon which is good. So, the runner has just counted out some swabs and she's documented those.

Part 2:

[Option 1]

Participant So the advantages of telling the staff to focus and speed up is that er the patient would be under anaesthetic for less time so it would be safer for the patient in terms of their anaesthetic risk erm and it would certainly er maybe encourage them in terms of providing an incentive to get them home faster erm however I think it's probably not the best option because erm telling someone to focus is often not the best way to get them to focus erm and they might need some other er interventions and the er worst case scenario here would be that erm that I taking this quite aggressive approach would alienate the other staff and thereby losing their sort of complicity in trying to get everything done as quickly as possible whilst keeping the patient as safe as possible so I think by taking this option you are probably going to decrease patient safety.

So she's obviously a bit tired because she's having to stretch her neck and she's wandering about. Got some stretching happening over here. This has taken a bit longer than it should have done so perhaps she wasn't prepared wasn't concentrating as well. I think she's not prepared. This might be a product of poor communication. Distracting the team. Ooh hello are we frustrated? There's some folded arms and tired faces here.

[Option 2]

Participant So ignoring the fatigue is probably a urm what most people would do erm because it means that everything would probably be done a little bit more expediently erm and it is sometimes not helpful to point out what other people might think of as failings in themselves so such as exhibiting signs of fatigue erm which might disenfranchise people from the operation erm however I think it's naïve to think that fatigue doesn't affect every part of the operation so ignoring it is really erm yeah just a juvenile thing to do really and could probably be tackled in a different way. Erm what you don't want staff to think is that you are unaware of your own situation and erm and that might sort of decrease their confidence in your abilities in this operation and in the next ones that you do.

[Option 3]

Participant: So so far this is definitely the best option so what you're doing here is you're acknowledging that there is a situation where patient safety could potentially be put at risk in that we are at the end of the day and everybody is going to be naturally fatigued what you're doing in this situation is sort of erm you're not putting any blame on any particular team member or on any particular kind of situation you're just acknowledging that this is where we are Erm you're showing that you have situational awareness erm and that will sort of improve confidence of the team in you erm certainly if you ask them to sort of keep an eye on each other as well it promotes a sense of teamwork erm and means that people don't feel individually responsible erm which might give them undue pressure erm which might lead to some mistakes so so far this is definitely the best option. The negative to this is of course that it's going to take a little bit of time again erm I think probably the time the few minutes that this would take would outweigh erm any time that would be er lapsed if there were any issue with patient safety.

P7e experimental

She's asked the scrub nurse for something. Erm she's passing the blade without it being guarded again so she's not using a kidney dish at any of the proper equipment erm that's could be an error to harm the scrub nurse or the surgeon that she's passing it to. Erm the scrub nurse was writing something on the board and then she's been asked not the scrub nurse the runner's been writing something on the board and the scrub nurse has asked her for something else so she is finished writing half way through erm so that could be an error when they're doing the counts at the end. Erm that scrub nurse has just erm emptied something onto a tray not the scrub nurse the runner's just emptied something onto the scrub nurse's tray without her looking so she might not know where it is erm it can be a problem for counting at the end. Erm she's carried on writing on the board but I can't see whether that the thing she's passed before or if she's written down what she just asked the scrub nurse so if she doesn't that'll be could be an error when the scrub nurse is counting up on her final count. Erm she's thrown swabs out of her sterile environment without informing the runners so you'd normally count what you're passing out so that they can help you count at the end and make sure its all correct. And again. Erm this it's been two different runners that's been sorting out the swabs without consulting the scrub nurse so she isn't aware of what they're doing erm so that could account for lost swabs or incorrectly counted swabs. Erm she the runner is writing something else on the board but there's no effective communication they're not communicating between them to make a plan up or describe to each other what they're doing. Erm I can't see an anaesthetist at all. Erm they're meant to count the tray before anything's taken out of it which they haven't done and something's also fallen on the floor which could also mean another error and they don't know what was actually on the tray to begin with because they haven't counted it.

Participant 8 during condition assessment

So from explaining situational awareness then if I was to be in surgery then yeah I agree you do scan the room to see what's happening, a and e people have it bad because they always have new cases so they always have to assess what is happening but if you're doing the same surgeries a lot of it is the same. So yeah, I see who's in the room and where in the surgery they are, like if I walked into a room I'm not sure how id be able to predict things , like I can predict what part of the surgery is next because I know the stages of the surgery but in terms of patient safety , like the video said to predict you need to have mental models but I'm not sure how, I mean unless I'm there maybe I will see something that makes me thing something bad might happen.. its hard to think about it without actually been there. But yes, I can try to think ahead more.

So he didn't say it in the video but from left to right that's kind of easy to hard isn't it. Like you start off with rules, I get told not to touch things still if I'm not sterile, like basic rules. But over time you know more the expectations than the rules like if it's the last surgery and its all over then yeah you can touch anything as it's all done with. Not a good example but that's more analyzing the situation isn't it.

Again, it's hard to think how I think, unless I'm in there doing something. Its good to know and if I'm doing something next time ill be thinking, right what kind of decision making am I using now and is it the best.

P10:

Its good to know about the way you think because I understand what it means , like perception, I get that it's a basic understanding of who's doing what, but then as you know what goes on then you start to understand how things should be and what should happen. I didn't know about the ways people decide, like the strategies. But I suppose it's like learning to drive, you follow the rules of how to use the car and on the road but eventually you don't have to think about, and you can spot problems better. Communication is the most important but there is a lot going on all the time, so you do get bombarded so its hard to get better unless you attend a lot and someone good at instructing can show you how to be I work as a scrub nurse in theatre so I believe having awareness, situational awareness is very important, it's the little things that you might do in theatre from the start that may become a habit in every surgery. So, it's the little mistakes that you might make that eventually just may become a habit that could put any patient's life at risk. So learning about situational awareness its taught me a lot because I feel that I can be more aware of my surroundings, not just think about the patient's safety but think about my own safety, safety of the staff around me as I may be held accountable for my own actions.

Participant 14 control

P14:

I've heard of the 3 levels of situational awareness as we had a project on personal skills and I was watching some YouTube videos about it and I thought it was useful , this is similar in that there are 3 levels and with more experience you can think about things to an extent where you can plan for things that other people miss. There was something about bias but if forgot what it said, but I know your attention can be bias to things if you are looking for them like you might think the surgeon is a bit mean so if he doesn't talk you might think he doesn't like you but really he's just tired. Or some procedure you think is going well but you've forgot to account for something as you were too busy with getting it done or something like that. The decision making types were the most useful out of this and I know there are proactive and reactive people, our lecturer mentioned if youre new you follow the rules and kind of wait for people to ask you to do things but your abilities change so you can do things before they need doing like preparing equipment and been more proactive and assessing the situation more that just 'what you should do'. The circles were a good visual on shared situational awareness my only concern with that is its impossible to figure out what someone is thinking, I've done it before in theatre where I meant something else and they meant something else I think I'd need practice more because knowing is good but its hard to actually do when your doing a million things at once in surgery you just don’t do them.

Figure 12C: After Intervention

Participant 1 Post-Video Assessment Transcript

So I'm looking around more initially as I want to take everything in first. The anaesthetists are not paying attention, they might not have to for now but if they get really involved in their conversation then they might not be aware if the stats change or a bag is empty later. The scrub nurse has to ask the surgeon again what he said. I said the music was loud last time, but I can see now as well it affects the staff. I didn't realise there are, or, there is everyone talking who's not in the main group doing the surgery, so I didn't realise that's why its noisy too I just thought it was the music. She's threw a swab on there. Oh, so the runner is called over to have a look at what the surgeon is doing, so she might be new to the procedure or he might just want them to understand the stage they are at, like in the other scenario when you tell your team what you're doing so everyone knows the next steps. Anaesthetists still talking, actually if there are 2, are they are swapping over? I'm guessing they are anaesthetists but maybe one is a trainee, hard to tell without seeing their badges or talking to them, and obviously I didn't see the team brief at the start of the procedure.

P2 Post-Video Assessment Transcript

Participant:

So, from what I know now it's not so much what they are doing but also how and why they are doing these things. So, the anaesthetists are talking but why is because they might not have much to do just now, but in that case take it out of the room as it is adding to the theatre noise. Or if they can wait until after I'm sure the talk can wait.

I see the scrub nurse lean in to hear this time, so thinking about it the combination of individual chat and even quiet music means the scrub nurse can't hear the surgeon a few feet away. the surgeon is really good to pause and keep the trainees informed. There is quite a lot of them so making sure they all can understand the exact stage they are at is good and means he is thinking about the team not just getting the tasks done. In fact, as we have done decision making, how does he decide when to stop and talk, he must know the next bit is important, so he could also say to all the staff, like announce the next bit, so everyone just pays attention again. Stepping over the wires there. It seems everyone is focused on their own thing they don't see her step over them. Or at least they don't say anything, but they should. So I see the runner being side tracked, I saw her here last time but didn't click on she had not written the full equipment down. I mean it's obvious she hasn't so someone can ask, and she will say she's not done it. But yes, I suppose if there is some mix up this could lead to something I might also not see. She's dropped a tool, and another. Has the music gone up?, maybe its just me paying more attention to things but it sounds like its more distracting.

Participant 7e post condition measure

the blade is being passed without being guarded, but for her that might be just the way that she likes it she might not be bothered about the risk of harm I didn't pick up on the other person asking questions as a distraction erm I can see that why that would be a distraction but some surgeons are very willing and interactive with the trainees where others just completely concentrate on what they're doing. Yeah so I picked up on the same thing as her (the feedback) but she might have had the same problem with it not being counted at the end and not being realised.

I didn't pick up on her leaving equipment on the drapes which it just depends on the surgeon some of them like them there, just to be ready to be picked up and others don't. It depends on the surgeon, some of them will have, like when I did obstetrics, they'll have everything out and others will just have what they're using at that time but I think it's just surgeon's preference or scrub nurse how tidy they are. Erm I don't think its room for error as long as the scrub nurse knows where her instruments are. Erm her mentioning the tools under the swabs I think that's the same problem as long as the scrub nurse knows where they are then they're not lost. It's if she doesn't know where her instruments are that then there's a problem. but she passed another sharp that wasn't guarded at all to the surgeon erm. I think she recognises different stuff from a surgeon's perspective than someone that scrubs and runners, erm like with the crossing off the documentation at the end and the scrub nurse had the back had her back to the whiteboard erm so the runner, like for us we would do it as in you're watching what your runner's doing because it's your instrument it's your swabs and you're the one whose accountable for them so it shouldn't be the runner just crossing it off when she's counted it should be both of you together counting the swabs. So when we've got five we'll like count out 1, 2, 3, 4, 5 with the runner yeah there's always two people on all the documentation its two people that's got to sign them so I think that's just they might have just dropped into a lazy which it's the same at some of the trusts I've worked at that they just trust each other too much.

P8c Post assessment

Participant: they're not really paying attention they've got to ask again what they've asked. Another thing that I noticed was that the scrub nurse like full on like leant on the patient like leant over the patient and like full on like leaned on them and again erm normally when they're incubated sometimes they have a tube kind of running down the patient like the tube that they've got connected to the ET tube so if that was happening and they leant on that that could have occluded the breathing tube. So that's again patient safety he won't be able to breathe for a minu te.

I mean the surgeon will have asked her like probably just to see where they are in the procedure really because obviously it's one of those procedures where sometimes you can't fully see what's going on all the time so sometimes we do have to lean over but one thing that we can do is just kind of stand up keep our hands to our self and just lean over as much as possible. Another thing we can get is a step so that's another thing. You could easily ask for a step, step on it and then she'll just be able to look over rather than full on having to lean on the patient which again could cause injury.

I mean she's literally got her back to the procedure. So like normally if the procedure's there we will try and sit and just kind of have a conversation like we are like now we can see the procedure but she's literally got her back to them so she might have no idea what's going on at the moment if they say oh something's happen she'll turned round and not she might not know where they are might not know what's going on so might not, so may take a bit more time to realise what they need which again could delay whatever's going on.

So I just heard something fall on the floor and as soon as something falls on the floor we need to say so say if it's a scissor we just need to say scissor on the floor and then because that's happened to me a few times erm so we says scissor on the floor whatever it is the runner should come over with some gloves on and pick it up and put it somewhere safe and then write down somewhere like on the whiteboard scissor on the floor because that tray needs to be fully complete when it goes back so if they don't know where that scissor is or whatever fell on the floor if something fell on the floor then they need to know where it is. They don't know if it's in the patient and will need to look back in the patient potentially. They need to make sure if that makes sense.

She's just put a blade on the side like there. She's just put a blade on the trolley I mean some people do do that but really in safety terms she should have it safe in like a kidney dish or somewhere safe because if she tries to grab something from here she could easily nick her hand on that. And those blades are really sharp it's and like any the retractors are really sharp as well so really. A surgeon really full on ripped my glove when I was in first year and I had to change it didn't cut me thankfully but it has happened like so yeah they're really sharp the amount of times I've nearly ripped my glove on stuff, I nick stuff and they're really sharp so again that should really be not in terms of patient safety but more in terms of staff safety really erm like that should be kind of in a kidney dish or on like a sharps pad where it is going to be safe because that's really dangerous.

She's struggling to get over a wire I mean that could be the suction she can't really take the suction down because obviously it needs to be there, but there's a lot of wires going on there and like a lot on them will be to the anaesthetic machine or your gas pipe lines erm but that's really bad. We do tend to you do see sometimes when there's loads of wires they'll like gather them together and get like a massive piece of like tape and tape them down which is the safest way to do it but like she could have easily fallen over and gone into a piece of equipment gone in to the surgeon or anything which is really bad.

She should really tell them where she's walking like she should say oh I'm just behind you so like I was saying with that guy that was stood behind me didn't tell me where he was and I didn't know. So if the surgeon couldn't yeah so she could accidentally go into him. If the surgeon stands back he could desterilise himself.

Participant 10 post condition measure

So situational awareness the 2 anaestnetisis are just talking and now and again looking at the monitor so they could be more linked into what the others are doing. erm the scrub nurse looks like she's been asked for something and she doesn't know where it is. Erm she's done the same again and she's not quite sure where her instruments are and she's also got a swab in her hand which means she's not she's trying to multitask and do two things at once erm which could mean to her either passing the wrong instrument which can be a pain for the surgeon or her mislaying this swab or not counting it correctly.

The lady on the left who's just hopped over the suction. She should have ideally not been behind the suction; the suction should have been behind her. The suction shouldn't even be there, they should bring the suction closer to the patient so it's not a trip hazard for anyone, because that suction, you're able to move it. The lady who's just walked away from the table she touched her face and her hand, she shouldn't be doing that after she's been so close to the patient. The lady who walked all the way around she did the right thing. She could have gone clockwise around the table, but she actually went all the way around the table which is good because then she's not in the sterile field

Participant 14 post condition measure

The scrub nurse she looks like she's looking for things like she's looking back and forth erm she looks a little bit stressed because I know I'd look like that if I was rushing. And she also looks like she is looking for a piece of equipment but can't find it because she swung right around erm just having a quick graze over all of her equipment. She looks like she's a bit, ah there we are. She's also got something in her hand that she hasn't got rid of it's like she doesn't really know where to put it. Erm she's doing good in the sense that she's looking back because she's seeing what's going on erm but yeah. She looks like she's getting a bit close to the erm. That isn't a good thing. She shouldn't just shove equipment down on the erm table erm on the patient erm because that's how things will go missing. And in a high stress environment erm that could make her panic as well as the swab that's also there. And she is [sighs] I think it might be a swab in her hand erm also that could err get lost if she lets go of it or whatever. And there she was sure she'd put something down there when she'd just put things over here or like on here. And no-one here's really helping her so that could also cause stress because they're all talking and instead of one these like HCAs or ODPs could come and just stand over and if someone was just looking like I would want as a circulator I'd be stood there looking at where she'd put things down and going aw maybe it's here aw maybe it might is this what you're looking for and maybe asking her if she needs any help looking for something. Ah see she found it over there. That's annoying. Good communication though. She has she has turned around erm because asked to write something on the board she straight away did it which shows she is kind of listening. They know their equipment that's good.

Table 1

Final Comparative SAGAT Results Table showing the SAGAT overview for participant 1. Improvements in SAGAT scores were mostly instrument related. DI=Did not Identify. Green=more accurate in post-measure. Red=less accurate in post-measure. Yellow= no change in accuracy.

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	12	12	12
Question 2	3	3	2
Question 3	1	1	2
Question 4	2	3	3
Question 5	1	1	1
Question 6	3	2	2
Question 7	90	120	161
Question 8	face	face	Face
Question 9	DI	Orange	Orange
Question 10	Full	Full	All
Question 11	DI	80	79 or 80
Question 12	No	4	4
Question 13	Yes	Yes	Yes
Question 14	Yes	Yes	no
Question 15	No	no	yes
Question 16	N/A	N/A	runner
Question 17	DI	Swab	Used Swab
Question 18	DI	DI	Item not present
Question 19	No	No	No
Question 20	2	2	2
Question 21	DI	No	Trainee
Question 22	no	no	no
Question 23	No	No	no
Question 24	DI	DI	Yes
Question 25	Yes, minimal	Yes	Yes
Question 26	DI	DI	Not there
Question 27	End	End	end
Question 28	Relaxed surroundings	Relaxed surroundings	Correct Answer

Table 2

Participant 2 SAGAT scores. There were minimal changes quantitatively as most were initially correct. These cues may have been easier to identify and recall due to the participant's frequent verbalizations of environment scanning.

Participant 2 SAGAT Results	Before Intervention	After intervention	Correct Answer
Question 1	11	11	12
Question 2	3	3	2
Question 3	2	1	2
Question 4	2	2	3
Question 5	1	1	1
Question 6	0	2	2
Question 7	150	120	161
Question 8	Face	face	Face
Question 9	Orange	Orange	Orange
Question 10	Full	Full	All
Question 11	80	80	79 or 80
Question 12	No	5	4
Question 13	Yes	Yes	Yes
Question 14	No	no	no
Question 15	No	yes	yes
Question 16	runner	Runner	runner
Question 17	Used swab	Swab	Used Swab
Question 18	DI	DI	not present
Question 19	No	No	No
Question 20	2	0	2
Question 21	DI	trainee	Trainee
Question 22	no	no	no
Question 23	No	No	no
Question 24	DI	DI	Yes
Question 25	Yes	Yes	Yes
Question 26	DI	DI	Not there
Question 27	End	End	end
Question 28	Calm and everyone seems comfy not busy like start	Runners comfortable, crossed arms fatigue	

Table 3

Table showing the Participant 3 SAGAT overview. Improvements in SAGAT scores were a variety of cue types. DI=Did not Identify. Green=more accurate in post-measure. Red=less accurate in post-measure. Yellow= no change

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	8	12	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	2	3	3
Question 5	1	1	1
Question 6	2	1	2
Question 7	150	180	161
Question 8	Face	Face	Face
Question 9	NA	orange	Orange
Question 10	All	All	All
Question 11	NA	80	79 or 80
Question 12	none	none	4
Question 13	no	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	NA	Used swab	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	NA	2	2
Question 21	Trainee	trainee	Trainee
Question 22	no	no	no
Question 23	NA	no	no
Question 24	didn't see	no	Yes
Question 25	no	Yes, but quiet	Yes
Question 26	didn't see	didn't see	Not there
Question 27	NA	end	end
Question 28	NA	Everyone relaxed like everything is underway already and bored	

Table 4

Participant 4 SAGAT scores that showed the most amount of questions being modified when compared to the pre-condition SAGAT

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	13	11	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	2	3	3
Question 5	1	1	1
Question 6	3	1	2
Question 7	180	150	161
Question 8	Face	Face	Face
Question 9	NA	yellow	Orange
Question 10	All	partial	All
Question 11	79	NA	79 or 80
Question 12	none	3	4
Question 13	none	yes	Yes
Question 14	no	no	no
Question 15	Yes	no	yes
Question 16	runner	NA	runner
Question 17	swab	NA	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	3	0	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn't see	didn't see	Yes
Question 25	Yes	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	middle	end	end
Question 28	people not doing much, but surgeon busy	relaxed environment, apart from the scrub nurse	

Table 5

Participant 5 SAGAT scores. Their accuracy toward the correct answers increase more than any other participant, with few decreases in accuracy.

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	8	9	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	2	3	3
Question 5	2	1	1
Question 6	2	2	2
Question 7	150	180	161
Question 8	head	Face	Face
Question 9	NA	yellow	Orange
Question 10	All	All	All
Question 11	NA	79	79 or 80
Question 12	none	none	4
Question 13	no	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	NA	knife	Used Swab
Question 18	didn’t see	didn’t see	Item not present
Question 19	no	yes	No
Question 20	0	3	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn’t see	didn’t see	Yes
Question 25	yes, but quiet	no	Yes
Question 26	didn’t see	no	Not there
Question 27	don’t know	don’t know	end
Question 28	--	--	

Table 6

Participant 6 SAGAT scores showed 7 improvements towards the correct answers

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	10	12	12
Question 2	3	3	2
Question 3	2	2	2
Question 4	2	3	3
Question 5	1	1	1
Question 6	5	2	2
Question 7	130	200	161
Question 8	Face	Face	Face
Question 9	orange	orange	Orange
Question 10	All	All	All
Question 11	NA	79	79 or 80
Question 12	none	4	4
Question 13	Yes	Yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	Used swab	swab	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	2	2	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn't see	yes	Yes
Question 25	yes, but quiet	yes, but quiet	Yes
Question 26	didn't see	no	Not there
Question 27	end	end	end
Question 28	runners comfortable	runners comfortable	

Table 7:

Participant 7 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	11	13	12
Question 2	3	3	2
Question 3	0	2	2
Question 4	3	3	3
Question 5	0	1	1
Question 6	2	4	2
Question 7	180	120	161
Question 8	Face	Face	Face
Question 9	NA	orange	Orange
Question 10	All	All	All
Question 11	NA	79	79 or 80
Question 12	none	none	4
Question 13	no	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	NA	swab	Used Swab
Question 18	didn’t see	didn’t see	Item not present
Question 19	no	no	No
Question 20	2	4	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn’t see	didn’t see	Yes
Question 25	Yes	yes, but quiet	Yes
Question 26	didn’t see	didn’t see	Not there
Question 27	beginning	beginning	end
Question 28	Saw to cut	saw being prepared	

Table 8

Participant 8 SAGAT results

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	11	12	12
Question 2	2	3	2
Question 3	2	1	2
Question 4	2	3	2
Question 5	0	1	1
Question 6	1	1	2
Question 7	158	151	161
Question 8	neck	neck	Face
Question 9	NA	orange	Orange
Question 10	All	All	All
Question 11	79	79	79 or 80
Question 12	2	1	4
Question 13	no	no	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	swab	used swab	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	4	2	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	Yes	Yes	Yes
Question 25	no	no	Yes
Question 26	didn't see	no	Not there
Question 27	end	end	end
Question 28		sn with blade down	

Table 9: Participant 9 SAGAT results

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	13	12	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	4	3	3
Question 5	0	0	1
Question 6	3	3	2
Question 7	180	180	161
Question 8	Body	Face	Face
Question 9	NA	NA	Orange
Question 10	All	All	All
Question 11	NA	NA	79 or 80
Question 12	none	none	4
Question 13	yes	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	NA	NA	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	yes	No
Question 20	0	0	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	yes	no	no
Question 24	didn't see	yes	Yes
Question 25	yes	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	end	end	end
Question 28	Knife was asked for, so cutting still	Knife was asked for, so cutting still	

Table 10: Participant 10 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	13	11	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	3	3	3
Question 5	3	0	1
Question 6	2	2	2
Question 7	180	140	161
Question 8	Face	Face	Face
Question 9	NA	NA	Orange
Question 10	All	All	All
Question 11	NA	NA	79 or 80
Question 12	none	none	4
Question 13	yes	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	NA	swab	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	2	1	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	yes	yes	Yes
Question 25	no	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	end	end	end
Question 28	everyone talking and casual	everyone talking and casual	

Table 11: Participant 11 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	10	11	12
Question 2	2	2	2
Question 3	1	2	2
Question 4	4	4	3
Question 5	0	0	1
Question 6	2	2	2
Question 7	150	140	161
Question 8	Face	Face	Face
Question 9	NA	NA	Orange
Question 10	All	All	All
Question 11	79	79	79 or 80
Question 12	none	3	4
Question 13	no	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	Used swab	na	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	0	0	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn't see	didn't see	Yes
Question 25	yes	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	middle	middle	end
Question 28	not suturing or doing a count	not suturing or doing a count	

Table 12: Participant 12 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	12	10	12
Question 2	2	3	2
Question 3	2	2	2
Question 4	3	2	3
Question 5	0	1	1
Question 6	2	0	2
Question 7	180	150	161
Question 8	Face	Face	Face
Question 9	NA	NA	Orange
Question 10	All	All	All
Question 11	NA	NA	79 or 80
Question 12	none	none	4
Question 13	yes	yes	Yes
Question 14	no	no	no
Question 15	Yes	Yes	yes
Question 16	runner	runner	runner
Question 17	used swab	na	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	4	4	2
Question 21	didn't see	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	yes	yes	Yes
Question 25	yes	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	middle	middle	end
Question 28	waiting for surgeon to move on to next part	waiting for surgeon to move on to next part	

Table 13: Participant 13 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	12	12	12
Question 2	3	3	2
Question 3	2	1	2
Question 4	2	2	2
Question 5	1	1	1
Question 6	2	2	2
Question 7	200	180	161
Question 8	Face	Face	Face
Question 9	Yellow	Orange	Orange
Question 10	Partial	All	All
Question 11	79	NA	79 or 80
Question 12	none	none	4
Question 13	no	yes	Yes
Question 14	no	yes	no
Question 15	no	none	yes
Question 16	999	999	runner
Question 17	NA	NA	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	0	0	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	didn't see	Didn't see	Yes
Question 25	no	yes	Yes
Question 26	didn't see	didn't see	Not there
Question 27	end	end	end
Question 28	NA	People look like they want to be finished	

Table 14: Participant 14 SAGAT scores

SAGAT Results	Before Intervention	After Intervention	Correct Answer
Question 1	13	11	12
Question 2	3	3	2
Question 3	1	2	2
Question 4	3	3	3
Question 5	3	0	1
Question 6	2	2	2
Question 7	180	140	161
Question 8	Face	Face	Face
Question 9	NA	NA	Orange
Question 10	All	All	All
Question 11	NA	NA	79 or 80
Question 12	none	none	4
Question 13	yes	yes	Yes
Question 14	no	no	no
Question 15	no	yes	yes
Question 16	runner	runner	runner
Question 17	NA	swab	Used Swab
Question 18	didn't see	didn't see	Item not present
Question 19	no	no	No
Question 20	0	0	2
Question 21	trainee	trainee	Trainee
Question 22	no	no	no
Question 23	no	no	no
Question 24	yes	yes	Yes
Question 25	no	no	Yes
Question 26	didn't see	didn't see	Not there
Question 27	end	end	end
Question 28	People talking and casual	People talking and casual	

Figure 13: Friedman's Test for the NTS category of Leadership

Control group:

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
preL	7	1.0000	.81650	.00	2.00
postL	7	.4286	.53452	.00	1.00

Ranks

	Mean Rank
preL	1.71
postL	1.29

Test Statistics[a]

N	7
Chi-Square	3.000
df	1
Asymp. Sig.	.083

a. Friedman Test

Experimental group:

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
preL	7	.1429	.37796	.00	1.00
postL	7	.4286	.53452	.00	1.00

Ranks

	Mean Rank
preL	1.36
postL	1.64

Test Statistics[a]

N	7
Chi-Square	1.000
df	1
Asymp. Sig.	.317

a. Friedman Test

Figure 14A: Results for the 2x2 univariate repeated mixed measures ANOVA

Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Time * Condition	CSPE1	Sphericity Assumed	9.143	1	9.143	3.368	.091	.219	3.368	.393
		Greenhouse-Geisser	9.143	1.000	9.143	3.368	.091	.219	3.368	.393
		Huynh-Feldt	9.143	1.000	9.143	3.368	.091	.219	3.368	.393
		Lower-bound	9.143	1.000	9.143	3.368	.091	.219	3.368	.393
	CSPE2	Sphericity Assumed	1.286	1	1.286	2.077	.175	.148	2.077	.264
		Greenhouse-Geisser	1.286	1.000	1.286	2.077	.175	.148	2.077	.264
		Huynh-Feldt	1.286	1.000	1.286	2.077	.175	.148	2.077	.264
		Lower-bound	1.286	1.000	1.286	2.077	.175	.148	2.077	.264
	CSC1	Sphericity Assumed	.036	1	.036	.028	.870	.002	.028	.053
		Greenhouse-Geisser	.036	1.000	.036	.028	.870	.002	.028	.053
		Huynh-Feldt	.036	1.000	.036	.028	.870	.002	.028	.053
		Lower-bound	.036	1.000	.036	.028	.870	.002	.028	.053
	CSC2	Sphericity Assumed	6.036	1	6.036	1.719	.214	.125	1.719	.227
		Greenhouse-Geisser	6.036	1.000	6.036	1.719	.214	.125	1.719	.227
		Huynh-Feldt	6.036	1.000	6.036	1.719	.214	.125	1.719	.227
		Lower-bound	6.036	1.000	6.036	1.719	.214	.125	1.719	.227
	CSC3	Sphericity Assumed	2.286	1	2.286	2.595	.133	.178	2.595	.317
		Greenhouse-Geisser	2.286	1.000	2.286	2.595	.133	.178	2.595	.317
		Huynh-Feldt	2.286	1.000	2.286	2.595	.133	.178	2.595	.317
		Lower-bound	2.286	1.000	2.286	2.595	.133	.178	2.595	.317
	CSPR1	Sphericity Assumed	.893	1	.893	1.000	.337	.077	1.000	.152
		Greenhouse-Geisser	.893	1.000	.893	1.000	.337	.077	1.000	.152
		Huynh-Feldt	.893	1.000	.893	1.000	.337	.077	1.000	.152
		Lower-bound	.893	1.000	.893	1.000	.337	.077	1.000	.152
	CSPR2	Sphericity Assumed	.893	1	.893	1.000	.337	.077	1.000	.152
		Greenhouse-Geisser	.893	1.000	.893	1.000	.337	.077	1.000	.152
		Huynh-Feldt	.893	1.000	.893	1.000	.337	.077	1.000	.152
		Lower-bound	.893	1.000	.893	1.000	.337	.077	1.000	.152
	CT1	Sphericity Assumed	9.143	1	9.143	15.059	.002	.557	15.059	.945
		Greenhouse-Geisser	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
		Huynh-Feldt	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
		Lower-bound	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
	CDI1	Sphericity Assumed	.321	1	.321	.278	.607	.023	.278	.078
		Greenhouse-Geisser	.321	1.000	.321	.278	.607	.023	.278	.078
		Huynh-Feldt	.321	1.000	.321	.278	.607	.023	.278	.078
		Lower-bound	.321	1.000	.321	.278	.607	.023	.278	.078
	CDU1	Sphericity Assumed	.143	1	.143	.207	.657	.017	.207	.070
		Greenhouse-Geisser	.143	1.000	.143	.207	.657	.017	.207	.070
		Huynh-Feldt	.143	1.000	.143	.207	.657	.017	.207	.070
		Lower-bound	.143	1.000	.143	.207	.657	.017	.207	.070
	CDS1	Sphericity Assumed	1.286	1	1.286	1.479	.247	.110	1.479	.202
		Greenhouse-Geisser	1.286	1.000	1.286	1.479	.247	.110	1.479	.202
		Huynh-Feldt	1.286	1.000	1.286	1.479	.247	.110	1.479	.202
		Lower-bound	1.286	1.000	1.286	1.479	.247	.110	1.479	.202
	CDS2	Sphericity Assumed	.143	1	.143	2.000	.183	.143	2.000	.256
		Greenhouse-Geisser	.143	1.000	.143	2.000	.183	.143	2.000	.256
		Huynh-Feldt	.143	1.000	.143	2.000	.183	.143	2.000	.256
		Lower-bound	.143	1.000	.143	2.000	.183	.143	2.000	.256
	CDI2	Sphericity Assumed	.036	1	.036	1.000	.337	.077	1.000	.152
		Greenhouse-Geisser	.036	1.000	.036	1.000	.337	.077	1.000	.152
		Huynh-Feldt	.036	1.000	.036	1.000	.337	.077	1.000	.152
		Lower-bound	.036	1.000	.036	1.000	.337	.077	1.000	.152
	STA1	Sphericity Assumed	.036	1	.036	.075	.789	.006	.075	.057
		Greenhouse-Geisser	.036	1.000	.036	.075	.789	.006	.075	.057
		Huynh-Feldt	.036	1.000	.036	.075	.789	.006	.075	.057
		Lower-bound	.036	1.000	.036	.075	.789	.006	.075	.057
	STA2	Sphericity Assumed	.000	1	.000	.000	1.000	.000	.000	.050
		Greenhouse-Geisser	.000	1.000	.000	.000	1.000	.000	.000	.050
		Huynh-Feldt	.000	1.000	.000	.000	1.000	.000	.000	.050
		Lower-bound	.000	1.000	.000	.000	1.000	.000	.000	.050
	SSI1	Sphericity Assumed	4.321	1	4.321	7.562	.018	.387	7.562	.714
		Greenhouse-Geisser	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
		Huynh-Feldt	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
		Lower-bound	4.321	1.000	4.321	7.562	.018	.387	7.562	.714

SCE1	Sphericity Assumed	.571	1	.571	.960	.347	.074	.960	.147
	Greenhouse-Geisser	.571	1.000	.571	.960	.347	.074	.960	.147
	Huynh-Feldt	.571	1.000	.571	.960	.347	.074	.960	.147
	Lower-bound	.571	1.000	.571	.960	.347	.074	.960	.147
SCS1	Sphericity Assumed	.893	1	.893	5.769	.033	.325	5.769	.598
	Greenhouse-Geisser	.893	1.000	.893	5.769	.033	.325	5.769	.598
	Huynh-Feldt	.893	1.000	.893	5.769	.033	.325	5.769	.598
	Lower-bound	.893	1.000	.893	5.769	.033	.325	5.769	.598
SL1	Sphericity Assumed	.143	1	.143	.923	.356	.071	.923	.144
	Greenhouse-Geisser	.143	1.000	.143	.923	.356	.071	.923	.144
	Huynh-Feldt	.143	1.000	.143	.923	.356	.071	.923	.144
	Lower-bound	.143	1.000	.143	.923	.356	.071	.923	.144
SL2	Sphericity Assumed	.571	1	.571	3.000	.109	.200	3.000	.357
	Greenhouse-Geisser	.571	1.000	.571	3.000	.109	.200	3.000	.357
	Huynh-Feldt	.571	1.000	.571	3.000	.109	.200	3.000	.357
	Lower-bound	.571	1.000	.571	3.000	.109	.200	3.000	.357
SL3	Sphericity Assumed	.036	1	.036	.300	.594	.024	.300	.080
	Greenhouse-Geisser	.036	1.000	.036	.300	.594	.024	.300	.080
	Huynh-Feldt	.036	1.000	.036	.300	.594	.024	.300	.080
	Lower-bound	.036	1.000	.036	.300	.594	.024	.300	.080

Figure 14B: SSI1 results for 2x2 mixed measures ANOVA

Descriptive Statistics

	Condition	Mean	Std. Deviation	N
Pre_SSI1	Experimental	.29	.756	7
	Control	.57	.535	7
	Total	.43	.646	14
Post_SSI1	Experimental	1.86	1.215	7
	Control	.57	.535	7
	Total	1.21	1.122	14

Levene's Test of Equality of Error Variances[a]

		Levene Statistic	df1	df2	Sig.
Pre_SSI1	Based on Mean	.000	1	12	1.000
	Based on Median	.167	1	12	.690
	Based on Median and with adjusted df	.167	1	10.800	.691
	Based on trimmed mean	.059	1	12	.812
Post_SSI1	Based on Mean	.732	1	12	.409
	Based on Median	.480	1	12	.502
	Based on Median and with adjusted df	.480	1	9.446	.505
	Based on trimmed mean	.784	1	12	.393

Tests of Within-Subjects Effects

Measure: SSI1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
factor1	Sphericity Assumed	4.321	1	4.321	7.562	.018	.387	7.562	.714
	Greenhouse-Geisser	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
	Huynh-Feldt	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
	Lower-bound	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
factor1 * Condition	Sphericity Assumed	4.321	1	4.321	7.562	.018	.387	7.562	.714
	Greenhouse-Geisser	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
	Huynh-Feldt	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
	Lower-bound	4.321	1.000	4.321	7.562	.018	.387	7.562	.714
Error(factor1)	Sphericity Assumed	6.857	12	.571					
	Greenhouse-Geisser	6.857	12.000	.571					
	Huynh-Feldt	6.857	12.000	.571					
	Lower-bound	6.857	12.000	.571					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: SSI1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	18.893	1	18.893	25.597	.000	.681	25.597	.996
Condition	1.750	1	1.750	2.371	.150	.165	2.371	.294
Error	8.857	12	.738					

a. Computed using alpha = .05

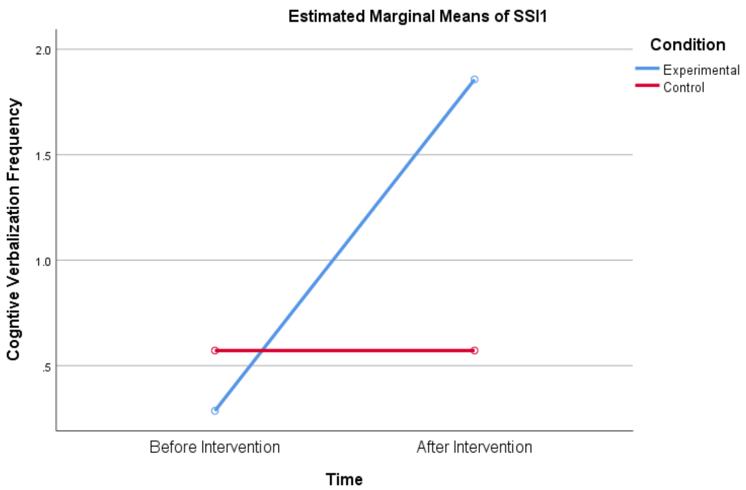


Figure 14C: SCS1 results for 2x2 mixed measures ANOVA

Tests of Within-Subjects Effects

Measure: SCS1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
factor1	Sphericity Assumed	2.286	1	2.286	19.200	.001	.615	19.200	.980
	Greenhouse-Geisser	2.286	1.000	2.286	19.200	.001	.615	19.200	.980
	Huynh-Feldt	2.286	1.000	2.286	19.200	.001	.615	19.200	.980
	Lower-bound	2.286	1.000	2.286	19.200	.001	.615	19.200	.980
factor1 * Condition	Sphericity Assumed	1.286	1	1.286	10.800	.007	.474	10.800	.854
	Greenhouse-Geisser	1.286	1.000	1.286	10.800	.007	.474	10.800	.854
	Huynh-Feldt	1.286	1.000	1.286	10.800	.007	.474	10.800	.854
	Lower-bound	1.286	1.000	1.286	10.800	.007	.474	10.800	.854
Error(factor1)	Sphericity Assumed	1.429	12	.119					
	Greenhouse-Geisser	1.429	12.000	.119					
	Huynh-Feldt	1.429	12.000	.119					
	Lower-bound	1.429	12.000	.119					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: SCS1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	2.286	1	2.286	19.200	.001	.615	19.200	.980
Condition	1.286	1	1.286	10.800	.007	.474	10.800	.854
Error	1.429	12	.119					

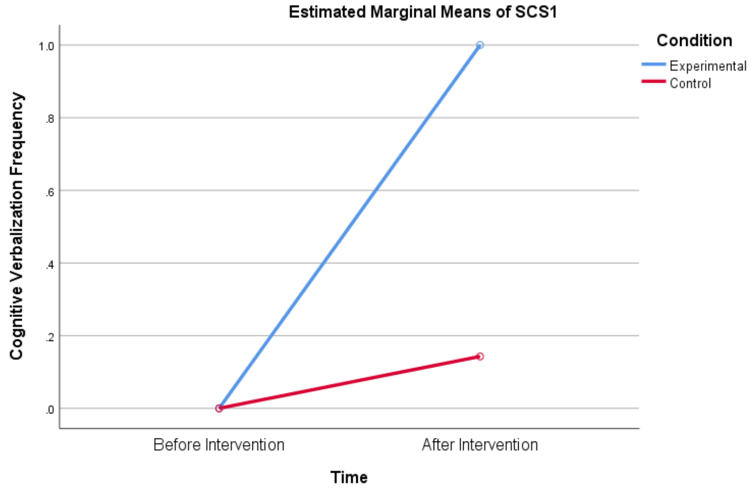


Figure 14D: CT1 results for 2x2 mixed measures ANOVA

Tests of Within-Subjects Effects

Measure: CT1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
factor1	Sphericity Assumed	11.571	1	11.571	19.059	.001	.614	19.059	.979
	Greenhouse-Geisser	11.571	1.000	11.571	19.059	.001	.614	19.059	.979
	Huynh-Feldt	11.571	1.000	11.571	19.059	.001	.614	19.059	.979
	Lower-bound	11.571	1.000	11.571	19.059	.001	.614	19.059	.979
factor1 * Condition	Sphericity Assumed	9.143	1	9.143	15.059	.002	.557	15.059	.945
	Greenhouse-Geisser	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
	Huynh-Feldt	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
	Lower-bound	9.143	1.000	9.143	15.059	.002	.557	15.059	.945
Error(factor1)	Sphericity Assumed	7.286	12	.607					
	Greenhouse-Geisser	7.286	12.000	.607					
	Huynh-Feldt	7.286	12.000	.607					
	Lower-bound	7.286	12.000	.607					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: CT1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	11.571	1	11.571	19.059	.001	.614	19.059	.979
Condition	9.143	1	9.143	15.059	.002	.557	15.059	.945
Error	7.286	12	.607					

a. Computed using alpha = .05

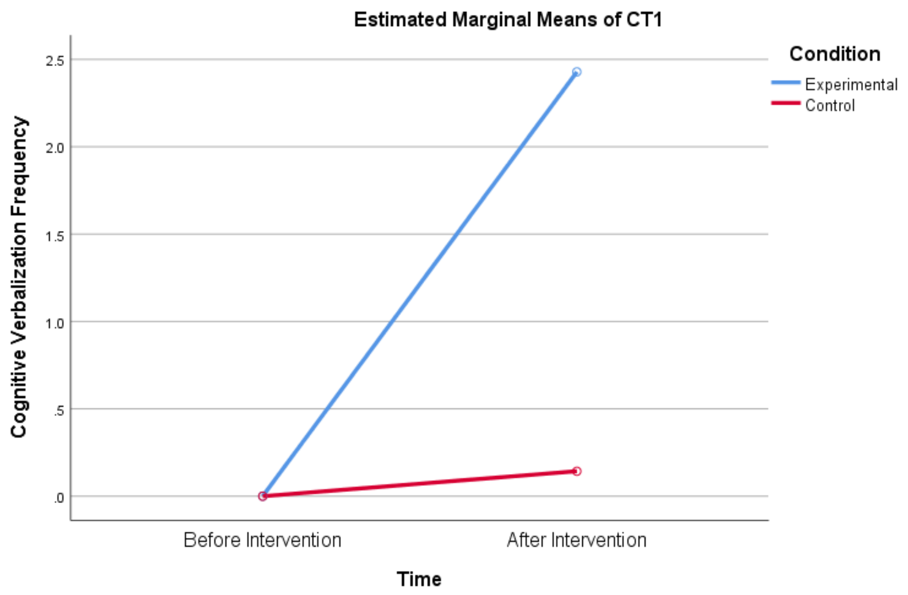


Figure 15: Participants' reflections after experiment

Participant 1:

Reflection after expert feedback SL2

It was Max-fax surgery wasn't it.

R: yes

Its common that you can put the... you have a bowl at the top where the surgeons with a drill, that you're not next to. And when she said they were assisting an anesthetist…

R: yes, it's a surgeon.

And to open stuff onto the tray without her noticing, it happens a lot. But normally someone points it out, so I didn't really think about it. There was a lot to take in, but normally if someone puts something onto your tray it's because you've asked for it. So, you would be like, where is it.

R: [explains a HF event in the video] did you see this at all?

Erm…. [doesn't respond, struggling to answer].. I didn't see the, well, I might have seen the actual behavior but I didn't know it was connected to something else. As in, I saw the other runner help with the used swabs but I didn't think that they might clump and look like 1, but she might not see it. I just saw her take over I didn't think what that means if the scrub nurse didn't separate the swabs right. So yeah if I saw that event before I feel I would have said she could count wrong, I just didn't see the first issue.

R: ok no worries let's stop.

*end

P2 Reflection typed v1 after feedback

'It was good to pause and assess things but I still missed some things. I wasn't that far off in terms of the number of things I identified, but there were a few links that I didn't see. I was going to write 'it's because I've only seen the video once' but obviously in real life things only happen once. I did justify in a similar manner to the feedback so I feel like I was thinking ahead good enough that if I was there, I would be able to say something or interject to make them aware of these things that's happening.

Participant: So I didn't see the bit where the scrub nurse nearly wrote something down where the runner nearly wrote something down but didn't erm and I think that's probably because I was distracted myself on the main event which was trying to see errors that were happening on the table rather than around the place and I think that's probably a product of the fact that when I am in theatre my priority is what's right in front of me and that I sort of have to very much tune out almost to all the external distractions coming from the runners and the other scrub nurses and things like that. Of which there are many [laughs]. So maybe that's a product of maybe why I sort of ignored that part of it.

The feedback showed me how much there is around me, but I still think, ok I still think to some extent i don't need to know every dynamic between everyone at every time, it's just too much but it has shown me how other factors can actually lead up to some impact on what I'm doing, not just small things like waiting for a tool for a few seconds longer, but it helps to show how my ability to keep an eye on the situation, if I can do it or learn to do it, can help me to lead others and keep control of things to steer the direction as we would like.

I don't like how I have to pause the video because by the time I've explained what I've seen then I do forget everything else that's going on, so it just takes me a few seconds to get back in to it. I feel that's why I've missed some things, as, even though it can pause it, there are so many non-technical things happening now I can specifically see what I'm looking for I think pausing and talking too much makes it harder to do. So, if I could just state quickly what I've seen and why I pointed it out, but then move on, would be easier to manage.'

P3 Written reflection

The vr is useful but I do feel a little dizzy. It's good for trainees but for practicing surgeons it might be limiting as they need to know more of the technical skills. Its more the team around them that its useful for. So from where the camera was it is good to see the theatre and you're able to pause it and I imagine rewind is useful. I did see some things the second time because the use of this kind of directs me to look at things, but it doesn't really tell you to. I just felt I was hyper alert its quite intense. But yeah, I didn't pause as much as I could so that might help. The decision-making section was the best, for me at least, as you can talk about the option, and then see how close you are. Not that I was wrong but just that second opinion from an established source helps you understand and feel more confident in your choice. Although in real life the options are not presented in front of you but I suppose you could ask the room what to do and them deliberate on it, based on this practice. I don't think i would use it unless I was a trainee. But yes, I did learn about non-technical skills even thought I felt I know them I was good to see them explained in front of you. I would suggest some feedback though during the scenarios as I did forget the points from the feedback as there were a lot of points

P5 written reflection

*participant was quiet during the last post assessment video. In non vr. I didn't say anything. I think she was just overloaded. I chose not to ask her to write anything here as I noticed she wasn't looking around much and felt it was right to finish the session.

P6 written reflection

P: It was good to be able to practice looking at the room and just pausing to guess what will be next. I was doing it a few times, trying to just say what is going to happen because of something I've seen. We have done it where we watch a video of a problem and it talks us thorough what is happening, but this was different as you have to say what the problems are. So, it made me think more why I was looking at something and what it means.

Participant 7 reflection

I do think it would be quite good maybe more erm maybe not when your practicing like not when you're qualified because obviously you are, you'll know it by that point but maybe more like in training maybe in your 1st and your 2nd years or even throughout for any form of profession really because some people will come into the course having no non-technical skills at all I mean I'm quite lucky where I have been in situations where I have been able to gain some and I've gained some more throughout this course so I do think it would be beneficial maybe depending on how it's done like how you've done it I mean you watch a video and then they you answer questions. I think that's quite a good thing because you can reflect on it, you can reflect on what you've seen and how you and you can learn from that and when you're in theatre you can you will it will help you maybe just think about it a bit more.